

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:59:01 ; Search time 4746.96 Seconds
(without alignments)
10471.464 Million cell updates/sec

Title: US-10-010-408-1
Perfect score: 1708
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GCCTAGATAAACACCCAAA 1708

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1621.6	94.9	1741	10	AF259981 Rattus no
2	1278	74.8	1734	6	AR210324 Sequence
3	1278	74.8	1734	6	AR210325 Sequence
4	1278	74.8	1734	10	AF100778 Mus muscu
5	1263.8	74.0	1739	10	AF126063 Mus muscu
6	901.2	52.8	137964	2	AC126895 Rattus no
7	881.8	51.6	226303	2	AC095418 Rattus no
8	645	37.8	61072	10	AL731698 Mouse DNA
9	645	37.8	216757	2	AL669906 Mus muscu
10	561.4	32.9	1266	6	AX076919 Sequence
11	561.4	32.9	1266	6	AX464186 Sequence
12	561.4	32.9	1283	9	AF083500 Homo sapi
13	561.4	32.9	1293	6	AR210322 Sequence
14	561.4	32.9	1293	6	AR210323 Sequence
15	561.4	32.9	1309	9	AF074604 Homo sapi
16	561.4	32.9	1427	9	AF100780 Homo sapi
17	561.4	32.9	1450	9	BC017782 Homo sapi
18	528.8	31.0	841	6	AR210338 Sequence
19	501.2	29.3	738	6	AR210337 Sequence
20	256.6	15.0	137964	2	AC126895 Rattus no
21	195.8	11.5	107260	9	AL139352 Human DNA
22	168	9.8	1585	5	NV1271167 Notophtha
23	163.8	9.6	2075	6	AR018957 Sequence
24	163.8	9.6	2075	6	AR118595 Sequence
25	163.8	9.6	2075	6	AR130891 Sequence
26	163.8	9.6	2075	6	AX206708 Sequence
27	163.8	9.6	2075	6	I11636 Sequence 1
28	163.8	9.6	2075	6	I32210 Sequence 1
29	163.8	9.6	2075	9	HUMCONGRO
30	163.8	9.6	2312	9	HSCCTGF
31	163.8	9.6	2998	6	AR119211 Sequence
32	163.8	9.6	2998	6	AR151276 Sequence
33	161	9.4	2267	6	AX206706 Sequence
34	161	9.4	2267	10	MUSFISP12B
35	161	9.4	2330	10	MUSTGFB
36	161	9.4	2334	10	BC006783
37	159.2	9.3	1496	4	SSU83916
38	158	9.3	2338	6	E37595
39	158	9.3	2338	10	AB023068
40	158	9.3	2345	6	AX402019
41	158	9.3	2345	10	AF120275
42	156.4	9.2	2350	6	AR194011
43	156	9.1	2190	5	AF463517
44	156	9.1	2288	5	GGA298335
45	154.8	9.1	2350	6	AR201286

ALIGNMENTS

RESULT 1
AF259981
LOCUS AF259981 1741 bp mRNA linear ROD 09-MAY-2000
DEFINITION Rattus norvegicus CCN family protein COP-1 (Cop-1) mRNA, complete cds.
ACCESSION AF259981
VERSION AF259981.1 GI:7739780
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1741)
AUTHORS Zhang,R., Averboukh,L., Zhu,W., Zhang,H., Jo,H., Dempsey,P.J.,

Coffey,R.J., Pardee,A.B. and Liang,P.
Identification of rCop-1, a new member of the CCN protein family,
as a negative regulator for cell transformation
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)
98414629
9742130
REFERENCE 2 (bases 1 to 1741)
AUTHORS Liang,P.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer
Center, 649 MRB II, Nashville, TN 37232, USA
FEATURES
source Location/Qualifiers
1..1741
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1..1741
/gene="Cop-1"
262..1014
/gene="Cop-1"
/note="secreted protein"
/codon_start=1
/product="CCN family protein COP-1"
/protein_id="AAF69011.1"
/db_xref="GI:7739781"
/translation="MRGSPILRLIATSFLCLLSMVCAQLCRTPCTPWTPPQCPQGV
LVLDGCGCKVCARRLTESCEHLHVCEPSQGLVCQPGAGPGGHGAVCLLDEDDGCEV
NGRRYLDGETFKPNCRVLCRDDGGFTCLPLCSEDEVLPWDPCPRKRIQVPGKCP
WVCDQVTPAIQRSAAQHQLSALVTPASADAPWPNWSTAWGPCSTTCGLGIATRVSN
QNRFCQLEIQRRCLPRCLAAARSHSSWNSAF"
BASE COUNT 386 a 491 c 480 g 384 t
ORIGIN
Query Match 94.9%; Score 1621.6; DB 10; Length 1741;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1680; Conservative 0; Mismatches 24; Indels 4; Gaps 4;
QY 1 GACGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGCCCTGGCAAGGCTGCAGCC 60
Dy 17 GACGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGG-CITGGCAAGGCTGCAG-C 74
QY 61 GCTGGGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 120
Dy 75 GCTGGGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 134
QY 121 CTGTCAGCTTGTCTAAAGCTTTAGCACTTGTGGTGGCTTGAAGCTGGCTTCACACACTGT 180
Dy 135 CTGTCAGCTTGTCTAAACCTTAGCACTTGTGGTGGCTTGAAGCTGGCTTCACACACTGT 194
QY 181 CACCTTCGTGGTGGCTCCACGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 240
Dy 195 CACCTTCGTGGTGGCTCCACGG-CTCACCTCCAGGTTTGAAGCTGGCTCCACAGGGAC 253
QY 241 ACGGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCTCT 300
Dy 254 ACGGTGACATGAGGGGCGAGCCCACTGATCCGCTTCTGGCCACTTCTCTCTCTCTCTCT 313
QY 301 TCTCAATGGTGTGTGCCAGCTGTGCCGACACACCCCTGTACCTGTCTTGGACACCCACC 360
Dy 314 TCTCAATGGTGTGTGCCAGCTGTGCCGACACACCCCTGTACCTGTCTTGGACACCCACC 373
QY 361 AGTGCCACACAGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTAAAGTGTGTGCAC 420
Dy 374 AGTGCCACACAGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTAAAGTGTGTGCAC 433
QY 421 GGAGGCTGGGGAGTCTTGGACACCTGTGATGTCTGCGACCCCGAGGCGCTGGTTT 480
Dy 434 GGAGGCTGACGGAGTCTTGGACACCTGTGATGTCTGCGAACCCCGAGGCGCTGGTTT 493
QY 481 GTCAGCCTGGGCGAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACG 540
Dy 494 GTCAGCCTGGGCGAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACG 553
QY 541 GTAGCTGTAGGTGAATGGCCGCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCA 600

554 GTGACTGTGAGGTGAATGGCCGACGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCA 613
QY 601 GGGTCCTGTGCGCTGTGATGACGGTGGCTTACCTGCCCTGCCCTGTGCAGTGAGGATG 660
Dy 614 GGGTCCTGTGCGCTGTGATGACGGTGGCTTACCTGCCCTGCCCTGTGCAGTGAGGATG 673
QY 661 TGGGGCTGCCAGCTGGGACTGCCCCACGCCCCCAAGAGAATACAGGTGCCAGGAAAGTCT 720
Dy 674 TGACGCTGCCAGCTGGGACTGCCCCACGCCCCCAAGAGAATACAGGTGCCAGGAAAGTCT 733
QY 721 GCCCGAGTGGGTATGTGACAGGGAGTGACACCGCGGATCCAGGGCTCCACGGCGCAAG 780
Dy 734 GCCCGAGTGGGTATGTGACAGGGAGTGACACCGCGGATCCAGGGCTCCGCGCGCAAG 793
QY 781 GACACCAACTTCTGCCCTTGTCACTCCTGCTGCTGTGATGCTCCTTGTCCAAATTTGA 840
Dy 794 GACACCAACTTCTGCCCTTGTCACTCCTGCTGCTGTGATGCTCCTTGTCCAAATTTGA 853
QY 841 GCACAGCCTGGCGCCCTGCTCAACCACTGTGGGTGGGCTGGGATAGCCACCCGAGTGTCCA 900
Dy 854 GCACAGCCTGGCGCCCTGCTCAACCACTGTGGGTGGGATAGCCACCCGAGTGTCCA 913
QY 901 ACCAGAACCGATCTGCCAACTGGAGATCCAAACGCCGCTGTGTCTGCCACAGACCCTGCC 960
Dy 914 ACCAGAACCGATCTGCCAACTGGAGATCCAAACGCCGCTGTGTCTGCCACAGACCCTGCC 973
QY 961 TGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCAACTGGGGATGCGGA 1020
Dy 974 TGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCAACTGGGGATGCGGA 1033
QY 1021 TACAGGCTGTCATCTCAGCAATGACCCCTAGGACAGGCCCTGGACTGCTGGTAGAT 1080
Dy 1034 TACAGGCTGTCATCTCAGCAATGACCCCTAGGACAGGCCCTGGACTGCTGGTAGAT 1093
QY 1081 GCTCTTCTCCAGTCTTGGCTGCAGTAACTGTCTGCTGGATTCACCTGTGTAGAGCC 1140
Dy 1094 GCTCTTCTCCAGTCTTGGCTGCAGTAACTGTCTGCTGGATTCACCTGTGTAGAGCC 1153
QY 1141 ACTGAGCGATCCTGCTGTGAGTAGGGGAGGAGCAGGTACAGCTCCAGTCTCTCTG 1200
Dy 1154 ACTGAGCGATCCTGCTGTGAGTAGGGGAGGAGCAGGTACAGCTCCAGTCTCTCTG 1213
QY 1201 GTTCAGCCTGGATTTCTGGGTCTCTGGCTCATTTCCCTCAAAACATCCCTGTACAAAAAG 1260
Dy 1214 GTTCAGCCTGGATTTCTGGGTCTCTGGCTCATTTCCCTCAAAACATCCCTGTACAAAAAG 1273
QY 1261 GACAACCAAAAAGACCTTTAAACCTAGGCTATACTGGGCAAAACCTGGCCACCCGTGCTGG 1320
Dy 1274 GACAACCAAAAAGACCTTTAAACCTAGGCTATACTGGGCAAAACCTGGCCACCCGTGCTGG 1333
QY 1321 GATAAGGTCAATGTTAGGACACAGCAGATTTGCCCTGAAACTTCCCAATTCCTTCTTGG 1380
Dy 1334 GATAAGGTCAATGTTAGGACACAGCAGATTTGCCCTGAAACTTCCCAATTCCTTCTTGG 1393
QY 1381 ACTTCTGTATGTTGTCCCCAAAGATGATGAATGAATCGTAAGTGTACCTTCCCTGACC 1440
Dy 1394 ACTTCTGTATGTTGTAAACCAAGATGATGAATGAATCGTAAGTGTACCTTCCCTGACC 1453
QY 1441 TGAGAACACCCCTGCTGCTCGGGAAGTATTCAGGGGCGAGAATTTCTGTGAACATGAAGA 1500
Dy 1454 TGAGAACACCCCTGCTGCTCGGGAAGTATTCAGGGGCGAGAATTTCTGTGAACATGAAGA 1513
QY 1501 GATGAATCACACTGTCTTAAAGAAATTCCTGAAAGTCCAGGAATTCAGCTTTGTATTTT 1560
Dy 1514 GATGAATCACACTGTCTTAAAGAAATTCCTGAAAGTCCAGGAATTCAGCTTTGTATT 1573
QY 1561 CAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTGGCAGGCC 1620
Dy 1574 CAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTAACA-GCC 1632
QY 1621 AGGCGCTTCTCTTTCAGCATGAGAAAGACAGGGACAGCAGAGTACTCTCTCTGGAGGA 1680

Db 1633 AGGGCCCTTCTTCAGCATGAGRAAGACAAGGGACAGCAGAGTACTCTCGTCTGGAGGA 1692

QY 1681 CTAGTCTAGCCTAGATAAACACCCAAA 1708

Db 1693 CTAGTCTAGCCTAGATAAACACCCAAA 1720

RESULT 2

AR210324

LOCUS AR210324 1734 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 17 from patent US 6387657.

ACCESSION AR210324

VERSION AR210324.1 GI:21512525

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1734)

AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.

TITLE WISP polypeptides and nucleic acids encoding same

JOURNAL Patent: US 6387657-A 17 14-MAY-2002;

FEATURES

source location/Qualifiers

1. .1734

/organism="unknown"

BASE COUNT 355 a 491 c 495 g 393 t

ORIGIN

Query Match 74.8%; Score 1278; DB 6; Length 1734;

Best Local Similarity 88.5%; Pred. No. 0;

Matches 1520; Conservative 0; Mismatches 165; Indels 33; Gaps 11;

QY 3 CGCTTCGATCTCCAGAGGACCCCTGGGTGGACAGGGGCCCTGGCAAGGCTGCAGCCGC 62

Db 13 CGCTCCTGATCTCCAGAGGACCCCGGCTGGACAGGGGCCCTGGCGAGGCTGCAGCTGC 72

QY 63 TG-GGCAGTGGCTTGGAAATGGAGGTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCC 121

Db 73 TGTGGCAGTAGCTTGGGATGGAGGTCTTCTTGCTGGGAACCTGAGGAGCTGAGAGGCTCC 132

QY 122 TGTGAC---CTTGTCCTAAAGTCTTAGCACTTGTGGTGGCTTGGCTTACACACTGTCA 178

Db 133 TGTGAGGCTCCTGTCTCTAAACTTTGGCACTTGGGTGGCTTGGGCTTACACACTGTCA 192

QY 179 GACACCTTCGTGGTGGCTCCACGGGCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGG 238

Db 193 GACACCTTCTTGGTGGCTCCTCGGC-----TCAGGTTTGAAGCTGGCTCCACAAGG 246

QY 239 ACACGGTGACATGAGGGGACGCCCACTGATCCATCTTCTGGCACTTCCCTTCCTCGCT 298

Db 247 ACACGGTGACATGAGGGGCAACCCCACTGATCCATCTTCTGGCACTTCCCTTCCTCGCAT 306

QY 299 TCTCTCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACTGTCTTGGACACCCACC 358

Db 307 TCTCTCAATGGTGTATTCCCAGCTGTGCCAGCACCCCTGTGCTGTCTTGGACACCCACC 366

QY 359 CCAGTGCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTAAAGTGTGTGC 418

Db 367 CCAGTGCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGTGCGAGTGTGTGC 426

QY 419 ACGGAGGCTGGGGAGTCTCGGACCACTGCTGCTGTGGATGGCTGTAAAGTGTGTGC 478

Db 427 ACGGAGGCTGGGGAGTCTCGGACCACTGCTGCTGTGGATGGCTGTGGCTGTGTGCGAGTGTGTGC 486

QY 479 TTGTCAGCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTTGGATGAGGATGA 538

Db 487 TTGTCAGCTGGGGCAGGCCCACTGGCCGCTGTGTGTGTGCTTCTCGAAGAGGATGA 546

QY 539 CGGTAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598

Db 547 CGGGAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGGAGACCTTTAAACCCCAATTG 606

QY 599 CAGGGTCCCTGTGCCGCTGTGATGACGGTGGCTTCACCTGCCTGCCGCTGTGCAGTGAGGA 658

Db 607 CAGGGTTTGTGCCGCTGTGATGACGGTGGTTTACCTGCCTGCCGCTGTGCAGTGAGGA 666

QY 659 TGTGCGGCTGCCAGCTGGGACTGCCACGCCCAAGAGATACAGGTGCCAGGAAGTG 718

Db 667 TGTGCGGCTGCCAGCTGGGACTGCCACGCCCAAGAGATACAGGTGCCAGGAAGTG 726

QY 719 CTGCCCCGAGTGGGTATGTGACCAGGGAGTGA--CACCGGCGATCCAGCGCTCCACGGC 775

Db 727 CTGCCCCGAGTGGGTGTGTACCAGGCAGTGATGCAGCCGGCAATCCAGCCCTCCTCAGC 786

QY 776 GCAAGGACACCAACTTTCTGCCCTTGTCACTCCTGCCTCTGCTGATGCTCCTTGTCCAAA 835

Db 787 CCAAGGACACCAACTTTCTGCCCTTGTCACTCCTGCATCTGCCGATGGCCCTGTCCAAA 846

QY 836 TTGGAGCACAGCCTGGGGCCCTGTCTCAACCACCTGTGGGCTGGGCATAGCACCCGAGT 895

Db 847 CTGGAGCACAGCCTGGGGCCCTGTCTCAACCACCTGTGGGTGGGCATAGCACCCGAGT 906

QY 896 GTCCAACCAGAACCGATTCTGCCAACTGGAGATCCAACGCCGCTGTGTCTGCCCAGACC 955

Db 907 ATCCAACCAGAACCGATTCTGCCAACTGGAGATCCAACGCTGCCTGTGTCTGCCAGACC 966

QY 956 CTGCCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCA-ACCTGGGGA 1014

Db 967 CTGCCTGGCATCCAGGAGCCACGGCTCATGGAACAGTGCCTTCTAGAGCCATTGCGGGGA 1026

QY 1015 TCGGATACAGGGCCTGCCATCTCTCAGCAAAATGACCCCTAGGACAGGCCCTGGACTGCTG 1074

Db 1027 TGTGGATACAGGGCCTGCCATCTCTCAGCAAAATGTCCCTAGGACAGGCCCTGGACTGATG 1086

QY 1075 GTAGATGCTCTTCTCCATGCTCTTGGCTGCAGTTAACTGTCTGCTTGGATTCACTGTGT 1134

Db 1087 GTAGATGCCCCCTCTCCATGCTCTTGGCTGCAGTTAACTGTCTGGGTGGATTTCAGTGTCC 1146

QY 1135 AGAGCACTGAGCGATCCCTGCTGTCTGAGGTAGGCGGAGCAGGTGACCACTCCAGT 1194

Db 1147 AGAGCCTCTGAGCGATCCCTGCTGTCTGAGGTGGGGAGCAGGTGACCACTCCATT 1206

QY 1195 TCTCTGG--TTCAGCCTGGAATCTCTGGGTCTCTCCTGGCTCATCTCTCAAAACATCCCTGT 1252

Db 1207 TCTCTGGATTCTGACCCAGGCTCTTGGGTCTCTCCTGGCTAGTCTCTCAAAACTTCCCTGT 1266

QY 1253 ACAAAGGACAAACCAAAAGACCTTTAAACCTAGGCTATAGTGGGCAACACCTGGCCACC 1312

Db 1267 ATGAAAAGGACAAACCAAAAGGACCTTTAAAGCTAAGCTGTAAGTGGGCAAGCCTGGCCACC 1326

QY 1313 GTGCTGGGGATAAGGTCAATGTTAGG-ACCAGACAGCAGATTGCCTGAAACTTCCAATTC 1371

Db 1327 ATGCTGGGGATAGTGACAGTAATAGGTACCAGGCAGCAGATGCTGAAACATCCAGGTC 1386

QY 1372 CCTTCTTGGACTTCTGTATGCTTGTCTCCCAAGATGATGAATAACTCGTAAGTGTACCT 1431

Db 1387 CCTTCTTGGACTTCTATGTGCTGT-CCCAAAGATTATGGTGACCTTGTAAAGTGTGCT 1445

QY 1432 TCCCTGACCTGAGAACACCCCTGCCTGCTCGGGAAGTATTTCAGGGGCAAGAATTCTCTGTGA 1491

Db 1446 TTCCTGATCTGAGAACACCCCTGCCCGGCTGGGA-----AGAATTTCTGGGA 1492

QY 1492 ACATGAAGAGAT-GAATCACACTGTCTCTTAAGCACTCGCAAAACAGGAGGCTCCACACCT 1550

Db 1493 ACATGAAGAGATGGAATCACACTATCTTAAAGACGCTTTCGAAGTCCAGGAACCTTGACC 1552

QY 1551 TTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAGGCTCCACACCT 1610

Db 1553 TTTGTATTTGTAAAAATACACATCTCTTAAATGCTCAACAAGCAAG-AGGCTCCACACTT 1611

QY 1611 CTGCAGGCCAGGGCCTTCTCTTTCAGCATGAGAAAGACAGGGACAGCAGAGTACTCTC 1670

Db 1612 CTGCAGGCCAGGGCCTTCTCTTTCAGCATGAGAGAGACAAGAAACAGTAGAGTACCCTC 1671

QY 1671 CTCTGAGGAGCTAGTCTAGCCTAGAAATAAACACCCAAA 1708

Db 1672 CTCTGGAGGACTGGCCCGGTCTGGAATAAACACCCAAA 1709

RESULT 3
AR210325/c
LOCUS AR210325 1734 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 18 from patent US 6387657.
ACCESSION AR210325
VERSION AR210325.1 GI:21512526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 18-MAY-2002;
FEATURES
 source 1..1734
 /organism="unknown"
BASE COUNT 393 a 495 c 491 g 355 t
ORIGIN

Query Match 74.8%; Score 1278; DB 6; Length 1734;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 1520; Conservative 0; Mismatches 165; Indels 33; Gaps 11;

QY 3 CGCTTCTGATCTCCAGAGGACCCTGGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGC 62
Db 1722 CGCTCCTGATCTCCAGAGGACCCTGGGCTGGGACAGGGGCCCTTGGCGAGGCTGCAGCTGC 1663

QY 63 TG-GGCAGTGGCTTGGAAATGGAGGCTCTTTATTACTGGGAACCTGAGGACCTAAGAGGCTCC 121
Db 1662 TGTGGCAGTAGCTTGGGATGGAGGCTCTTTCTGCTGGGAACCTGAGGAGCTGAGAGGCTCC 1603

QY 122 TGTGAG---CTTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCA 178
Db 1602 TGTGAGGCTCTTGTCTCTAAACTCTTGGCACTTGGGTGGCTTGGGCTTCACACACTGTCA 1543

QY 179 GACACCTTCGTGGTGGCTCCACGGCCCTACCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238
Db 1542 GACACCTTCTTGGTGGCTCTCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 1489

QY 239 ACACGGTGACATGAGGGGACGCCCACTGATCCATCTTCTGGCCACTTCCTCTGCCT 298
Db 1488 ACACGGTGACATGAGGGGCAACCCCACTGATCCATCTTCTGGCCACTTCCTCTGCAT 1429

QY 299 TCTCTCAATGGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCAACC 358
Db 1428 TCTCTCAATGGTGTATTCACAGCTGTGCCGACACCCCTGTGCCTGTCTTGGACACCAACC 1369

QY 359 CCAGTGCCCAACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGCTGTAAAGTGTGTGC 418
Db 1368 CCAGTGCCCAACCGGGGTACCCCTGGTGTGGATGGCTGTGGCTGCTGTGAGTGTGTGC 1309

QY 419 ACGGAGGCTGGGGAGTCCCTGCGACCACTGTCTGCGACCCCGCCAGGGCCCTGGT 478
Db 1308 ACGGAGGCTGGGGAGTCCCTGCGACCACTGTCTGCGACCCCGCCAGGGCCCTGGT 1249

QY 479 TTGTAGCCTGGGCGAGGCCCTGGCGGCGCATGGGGCTGTGTCTCTTGGATGAGGATGA 538
Db 1248 TTGTAGCCTGGGCGAGGCCCTGAGTGGCCGTGGTGTGTGTGCTCTTCGAAGAGGATGA 1189

QY 539 CGGTAGCTGTAGGTGAATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAATTG 598
Db 1188 CGGGAGCTGTAGGTGAATGGCGCAGGTACCTGGATGGGAGACCTTTAAACCCAATTG 1129

QY 599 CAGGGTCTGTGCGGCTGTGATGACGGTGGCTTACCTGCCTGCCCTGTGTCAGTGAGGA 658
Db 1128 CAGGGTCTTGTGCGGCTGTGATGACGGTGGTTCACCTGCCTGCCCTGTGTCAGTGAGGA 1069

QY 659 TGTGCGGCTGCCAGCTGGGACTGCCCCACGCCCAAGAGAAATACAGGTGCCAGGAAGTG 718

Db 1068 TGTGCGGCTGCCCAGCTGGGACTGCCACGCCCCAGGAGAAATACAGTGCCAGGAAGTG 1009

QY 719 CTGCCCCGAGTGGGTATGTGACCAGGGAGTGA---CACCGGGGATCCAGCGCTCCACGGC 775
Db 1008 CTGCCCCGAGTGGGTGTGTGACCAGGCACTGATGCAGCCGGCAATCCAGCCCTCCTCAGC 949

QY 776 GCAAGGACACAACTTTCTGCCCTTGTCACTCCTGCCTCTGCTGCTGCTGCTGCTGCTGCTCAA 835
Db 948 CCAAGGACACAACTTTCTGCCCTTGTCACTCCTGCATCTGCCGATGGCCCTGTGCTCAA 889

QY 836 TTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGTGGCTGGGCATAGCCACCCGAGT 895
Db 888 CTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGTGGCTGGGCATAGCCACCCGAGT 829

QY 896 GTCCAACCAAGCCGATTTCTGCCAACTGGAGATCCAACGCGCCCTGTGTCTGCTGCTGCTGCTG 955
Db 828 ATCCAACCAAGCCGATTTCTGCCAACTGGAGATCCAACGCGCTGCTGCTGCTGCTGCTGCTG 769

QY 956 CTGCCCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCA-ACTGGGA 1014
Db 768 CTGCCCTGGCAGCCAGGAGCCACGCTCATGGAACAGTGTCTTCTAGAGCCATTGCGGGA 709

QY 1015 TGGGATACAGCGCCCTGCCATCTCAGCAATGACCCCTAGGACCAAGCCCTGGACTGCTG 1074
Db 708 TGTGGATACAGCGCCCTGCCATTCTCAGCAATGTCCTTAGGACCAAGCCCTGGACTGATG 649

QY 1075 GTAGATGCTCTTCTCCATGCTCTTGGCTGCACTTAAGTGTCTGCTGCTGCTGCTGCTGCTGCT 1134
Db 648 GTAGATGCTCTTCTCCATGCTCTTGGCTGCACTTAAGTGTCTGCTGCTGCTGCTGCTGCTGCT 589

QY 1135 AGAGCCACTGAGCGGATCCCTGCTGCTGCTGAGGTAGGCGGAGCAGGTGACCACTCCAGT 1194
Db 588 AGAGCCCTGAGCGGATCCCTGCTGCTGCTGAGGTGGGGAAGCAGGTGACCACTCCATT 529

QY 1195 TCTCTGG--TTGAGCCTGGAATTTCTGGGTCTCTGGCTCATTTCTCAAAAACATCCCTGT 1252
Db 528 TCTCTGGATTCTGACCCAGGCTTCTGGGTCTCTGGCTAGTTCTCTCAAAAACATCCCTGT 469

QY 1253 ACAAAAAGGACAAACCAAAAGACCTTTAAACCTAGGCTATACTGSSCAAAACCTGGCCACC 1312
Db 468 ATGAAAAGGACAAACCAAAAGACCTTTAAAGCTAAGCTGTACTGSSCAAGCCTGGCCACC 409

QY 1313 GTGCTGGGATAGGTCAATGTTAGG-ACCAGACAGCAGATTGCCCTGAAACTTCCAATTTC 1371
Db 408 ATGCTGGGATAGTGACAGTAATAGGTACCAGGCAGCAGATTGCCCTGAAACATCCAGGTC 349

QY 1372 CCTTCTTGGACTTCTGTATGCTTGTCCCAAAAGATGATGAATGAACCTGTAAGTGTACCT 1431
Db 348 CCTTCTTGGACTTCTATGTGCTGT-CCCAAGATTATGGGTGACCTTGTAAAGTGTGCCT 290

QY 1432 TCCCTGACCTGAAACACCCCTGCCTGCTGCGGGAAGTATTACAGGGGCAAGAAATCTCTGTGA 1491
Db 289 TTCTGTATCTGAGAAACACCCCTGCCCGGCTGGGA-----AGAAATTTCTGGGA 243

QY 1492 ACATGAAGAGATGAATCACACTGTCTTAAAGAAATTCCTGAAAGTCCAGGAACCTTGAGC 1550
Db 242 ACATGAAGAGATGAATCACACTTATCTTAAAGAGCGTTTGCCAAAGTCCAGGAACCTTGACC 183

QY 1551 TTTGTATTTTCAAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCT 1610
Db 182 TTTGTATTTTGAATAATACACATCTCTTAAATGCTCACAAGCAAG-AGGCTCCACACTT 124

QY 1611 CTGGCAGGCCAGGCGCTTCTCTTTCAGATGAGAAAGACAAGGACAGCAGAGTACTCTC 1670
Db 123 CTGGCAGGCCAGGCGCTTCTCTTTCAGATGAGAGAGACAAGGAACAGTAGAGTACCCTC 64

QY 1671 CTCTGGAGGACTAGTCTAGCCTAGATAATAAACACCCAAA 1708
Db 63 CTCTGGAGGACTGGCCCGGTCTGGAATAAACACCCAAA 26

AF100778		linear	ROD 17-DEC-1998
LOCUS	AF100778	1734 bp	mRNA
DEFINITION	Mus musculus connective tissue growth factor related protein WISP-2 (Wisp2) mRNA, complete cds.		
ACCESSION	AF100778		
VERSION	AF100778.1	GI:4028578	
KEYWORDS			
SOURCE			
ORGANISM	Mus musculus. Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1734)		
AUTHORS	Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A., Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C., Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D., Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.		
TITLE	WISP genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)		
MEDLINE	99061933		
PUBMED	9843955		
REFERENCE	2 (bases 1 to 1734) Pennica,D. Direct Submission		
AUTHORS	Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA		
TITLE	Location/Qualifiers		
JOURNAL	1..1734		
FEATURES	/organism="Mus musculus" /db_xref="taxon:10090" /cell_line="C57MG" /cell_type="epithelial" /tissue_type="mammary" /note="transformed by Wnt-1"		
source	1..1734		
gene	/gene="Wisp2" 257..1012		
CDS	/gene="Wisp2" /codon_start=1 /product="connective tissue growth factor related protein WISP-2" /protein_id="AAC96320.1" /db_xref="GI:4028579" /translation="MRGNPLIHLAISFLCILSMVYSQLCPAPCAPWTPTPCPPGV LVLDGCGCRVCARRLGSCDHIVCDPSQGLVCQPAGSPGRGAICLFEDDGSCEV NGRYLDGETFKPNCRVLCRDCDGFTCLPLCEDSVRLPSWDGPRPRIQVPGRCCPE WVCDQAVMQPAIQPSAQQHQLSALVT PASADGPCPNWSTAWGPCSTTCGLGIATRS NONRFQLEIQRRLLCSRPCLASRHSGWSNSAF"		
BASE COUNT	355 a 491 c 495 g 393 t		
ORIGIN			
	Query Match 74.8%; Score 1278; DB 10; Length 1734;		
	Best Local Similarity 88.5%; Pred. No. 0;		
	Matches 1520; Conservative 0; Mismatches 165; Indels 33; Gaps 11;		
QY	3 CGCTTGTGATCTCCAGAGACCCTGGGGTGGCACAGGGGCCCTTGCCAAGGCTGCAGCCGC 62		
Dg			
	13 CGCTCCTGATCTCCAGAGGACCCCCGGCTGGCACAGGGGCCCTTGCCGAGGCTGCAGCTGC 72		
QY	63 TG-GGCAGTGGCTTGGAAATGGAGGTTATTACTGGGAAC T G A G G A C T G A G G A G G C T C C 121		
Dg			
	73 TGTGGCAGTAGCTGGGATGGAGGTCCTTCTCTGCTGGGAAC T G A G G A C T G A G G A G G C T C C 132		
QY	122 TGTCA G - - - CTGTCTCAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCA 178		
Dg			
	133 TGTCAAGGCTCTGTCTCTAAACTCTTGGCACTTGGGTGGCTTGGGCTTCACACACTGTCA 192		
QY	179 GACACCTTCGTGGTGGCCTCCACGGCCCTCACCTTCACGTTTGAAGCTGGCTCCACAAGGG 238		
Dg			
	193 GACACCTTC GTGGTGGCCT CCTCGGCC - - - - - TCAGGTTTGAAGCTGGCTCCACAAGGG 246		
QY	239 ACACGGTGACATGAGGGGCAGCCCACCTCATTCATCTTCTGGCCACTTCCCTCTCTGCCT 298		

QY	839	GAGCACAGCCTGGGGCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGTGC	898
Db	835		
QY	899	CAACAGAACCGATTCTGCCAACTGGAGATCCAAACGCCGCTGTGTCTGCCAGACCCCTG	958
Db	895		
QY	959	CCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTAAGGCCA-AC	1017
Db	955		
QY	1018	GGATACAGGCCTGCCATCCTCAGCAAAATGACCCCTAGGACGAGCCCTGGACTGCTGGTA	1077
Db	1015		
QY	1078	GATGCTCTTCCATGCTCTTGGCTGCAGTTAACTGTCCTGTGGATTCACTGTGTAGA	1137
Db	1075		
QY	1138	GCCACTGAGCGATCCCCTGCTGTCTGAGGTAGGCGGAGCAGGTGACCAAGCTCCAGTTCT	1197
Db	1135		
QY	1198	CTGG--TTCAGCCTGGAATCTCGGTTCTCCTGGCTCAFTCCTCAAAACATCCCTGTACA	1255
Db	1195		
QY	1256	AAAAGGACAACCAAAAGACCTTTAAACCTAGGCTATAGTGGGCAAAACCTGGCCACCGTG	1315
Db	1255		
QY	1316	CTGGGGATAAGGTCAATGTTAGG-ACCAGACAGCAGATTGCCTGAAACTTCCATTCCT	1374
Db	1315		
QY	1375	TCCTGGACTTCTGTATGCTTGTCCCCAAAGATGATGAATGAACCTCGTAAGTGTA	1434
Db	1375		
QY	1435	CTGACCTGAGAACACCCCTGCTCGGGAAGTATTCAGGGGCAGAAATCTCTGTGAACA	1494
Db	1434		
QY	1495	TGAAGAGAT-GAATCACACTGCTCCTTAAGAAATTCCTGAAAGTCCAGGAAC	1553
Db	1481		
QY	1554	GTAATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTG	1613
Db	1541		
QY	1614	GCAGGCCAGGCCTTTCTCTTCAGCATGAGAAAGACACAGGACAGCAGAGTACTCTCCTC	1673
Db	1600	GCAGGCCAGGCCTTTCTCTTCAGCATGAGAGAGACACAGGAACAGTAGAGTACCCCTCT	1659
QY	1674	TGAGGAGCTAGTCTAGCCTAGATAATAACACCCAAA	1708
Db	1660		
		TGGAGGACTGGCCCGGTTTGGAAATAACACCCAAA	1694

RESULT 6
AC126895
LOCUS AC126895 137964 bp DNA linear HTG 24-JUL-2002
DEFINITION Rattus norvegicus clone CH230-301E4, *** SEQUENCING IN PROGRESS
***, 49 unordered pieces.
ACCESSION AC126895
VERSION AC126895.1 GI:21724040
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.
1 (bases 1 to 137964)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmami,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.
Unpublished
Direct Submission
2 (bases 1 to 137964)
Worley,K.C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 137964)
Worley,K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G2HG
Center clone name: CH230-301E4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89474 bases at least Q40
Consensus quality: 93422 bases at least Q30
Consensus quality: 96506 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

QY	1195	TCTCTGGTT	CAGCCTGGAATTCTGGGTTCTCCTGGCTCATTCCTCAAAACATCCCTGTAC	1254
Db	44481	TCTCTGGTT	CAGCCTGGGATTCTGGGTTCTCCTGGCTCATTCCTCGGAACATCCCTGTAC	44540
QY	1255	AAAAAGGACAACCA	AAAAAGACCTTTAAACCTAGGCTATACTGGGCAAAACCTGGCCACCGT	1314
Db	44541	AAAAAGGACAACCA	AAAAAGACCTTTAAACCTAGGCTATACTGGGCAAAACCTGGCCACCGT	44600
QY	1315	GCTGGGATAAGGT	CAATGTTAGGACCAGACAGCAGATTCGCTGAAACTTCCAATCCCT	1374
Db	44601	GCTGGGATAAGGT	C-TTGTTAGGACCAGACAGCAGATTCGCTGAAACTTCCAATCCCT	44659
QY	1375	TCTTGGACTTCT	GTATGCTTGTCOCCAAAGATGATGAATGAACCTCGTAAGTGTACCTTC	1434
Db	44660	TCTTGGACTTCT	GTATGCTTGTCOCCAAAGATGATGAATGAACCTCGTAAGTGTACCTTC	44719
QY	1435	CTGACCTGAGAA	CACCCCTGCCTCGGGGAAGTATTCAGGGGCAGAAATTCCTGTGAACA	1494
Db	44720	CTGACCTGAGAA	CACCCCTGCCTCGGGGAAGTATTCAGGGGCAGAAATTCCTGTGAACA	44779
QY	1495	TGAAGAGATGA	ATCACACTGTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTG	1554
Db	44780	TGAAGAGATGA	ATCACACTGTCTTAAGAAATTCCTCAAGTCCAGGAACCTTGAGCTTG	44839
QY	1555	TATTTTCAGGA	ATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTGG	1614
Db	44840	TATTTTCAGGA	ATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTGG	44899
QY	1615	CAGGCAGGGCCT	TTTCTCTTCAGCATGAGAAAGACAAAGGGACAGAGTACTCTCCTCT	1674
Db	44900	TAGGCCAGGGCCT	TTTCTCTTCAGCATGAGAAAGACAAAGGGACAGAGTACTCTCCTCT	44959
QY	1675	GGAGGACTAGT	CTAGCCTAGATAAAACACCCAAA	1708
Db	44960	GGAGGACTAGT	CCAGCCTAGATAAAACACCCAAA	44993

RESULT 7	AC095418	226303 bp	DNA	linear	HTG 11-JUL-2002
LOCUS	Rattus norvegicus clone CH230-7C10, *** SEQUENCING IN PROGRESS ***	51 unordered pieces.			
DEFINITION	AC095418	GI:21717893			
ACCESSION	AC095418.3	GI:21717893			
VERSION	HTG; HTGS_PHASE1.				
KEYWORDS	Norway rat.				
SOURCE	Rattus norvegicus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 226303)				
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,				

Lozado,R.J.,	Lu,X.,	Lucier,A.,	Lucier,R.,	Luna,R.,	Ma,J.J.,
Mareshwari,M.,	Mapua,P.,	Martin,R.,	Martindale,A.,	Martinez,E.,	Massey,E.,
Mawhiney,E.,	McLeod,M.P.,	Meador,M.,	Mei,G.,	Metzker,M.,	Miner,G.,
Miner,Z.,	Mitchell,T.,	Mohabbat,K.,	Morgan,M.,	Morris,S.,	Moser,M.,
Neal,D.,	Newton,J.,	Newton,N.,	Nguyen,A.,	Nguyen,N.,	Nguyen,N.,
Nickerson,E.,	Nwokenkwo,S.,	Oguh,M.,	Okwuonu,G.,	Oragunye,N.,	Oviedo,R.,
Pace,A.,	Payton,B.,	Peery,J.,	Perez,L.,	Peters,L.,	Pickens,R.,
Primus,E.,	Pu,L.L.,	Quiles,M.,	Ren,Y.,	Rives,M.,	Rojas,A.,
Rojobokan,I.,	Rolfe,M.,	Ruiz,S.,	Savery,G.,	Scherer,S.,	Scott,G.,
Shen,H.,	Shooshtari,N.,	Sisson,I.,	Sodergren,E.,	Sonaik,T.,	Sparks,A.,
Stanley,H.,	Stone,H.,	Sutton,A.,	Svatek,A.,	Tabor,P.,	Tamerisa,K.,
Tang,H.,	Tansey,J.,	Taylor,C.,	Taylor,T.,	Telfrod,B.,	Thomas,N.,
Thomas,S.,	Usmani,K.,	Vasquez,L.,	Vera,V.,	Villalon,D.,	Vinson,R.,
Wang,Q.,	Wang,S.,	Ward-Moore,S.,	Warren,R.,	Washington,C.,	Watlington,S.,
Williams,G.,	Williamson,A.,	Wleczyk,R.,	Wooden,S.,	Worley,K.,	Wu,C.,
Wu,Y.,	Wu,Y.F.,	Zhou,J.,	Zorrilla,S.,	Nelson,D.,	Weinstock,G.
Gibbs,R.	Direct Submission	Unpublished	2 (bases 1 to 226303)	Worley,K.C.	Direct Submission
Submitted (17-SEP-2001)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	3 (bases 1 to 226303)	Worley,K.C.	Direct Submission	Submitted (11-JUL-2002)
Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	On Jul 10, 2002 this sequence version replaced gi:17941885.	----- Genome Center	Center: Baylor College of Medicine	Center code: BCM	Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu	----- Project Information	Center project name: GCDF	Center clone name: CH230-7C10	----- Summary Statistics	Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads	Assembly program: Phrap; version 0.990329	Consensus quality: 185950 bases at least Q40	Consensus quality: 190362 bases at least Q30	Consensus quality: 193076 bases at least Q20	-----
* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).	* NOTE: This is a 'working draft' sequence. It currently consists of 51 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.	* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	1 1051: contig of 1051 bp in length	1052 1151: gap of unknown length	1152 2286: contig of 1135 bp in length
2287 2386: gap of unknown length	2387 3433: contig of 1047 bp in length	3434 3533: gap of unknown length	3534 4614: contig of 1081 bp in length	4615 4715 6581: contig of 1867 bp in length	6582 6682 7773: contig of 1092 bp in length
7774 7874 9530: contig of 1657 bp in length	9531 9630: gap of unknown length				

* 9631 10661: contig of 1031 bp in length
* 10662 10761: gap of unknown length
* 10762 11958: contig of 1197 bp in length
* 11959 12058: gap of unknown length
* 12059 13140: contig of 1082 bp in length
* 13141 13240: gap of unknown length
* 13241 14777: contig of 1537 bp in length
* 14778 14877: gap of unknown length
* 14878 16063: contig of 1186 bp in length
* 16064 16163: gap of unknown length
* 16164 17585: contig of 1422 bp in length
* 17586 17685: gap of unknown length
* 17686 18916: contig of 1231 bp in length
* 18917 19016: gap of unknown length
* 19017 20305: contig of 1289 bp in length
* 20306 20405: gap of unknown length
* 20406 21537: contig of 1132 bp in length
* 21538 21637: gap of unknown length
* 21638 23703: contig of 2066 bp in length
* 23704 23803: gap of unknown length
* 23804 26306: contig of 2503 bp in length
* 26307 26406: gap of unknown length
* 26407 28431: contig of 2025 bp in length
* 28432 28531: gap of unknown length
* 28532 30615: contig of 2084 bp in length
* 30616 30715: gap of unknown length
* 30716 33215: contig of 2500 bp in length
* 33216 33315: gap of unknown length
* 33316 35357: contig of 2042 bp in length
* 35358 35457: gap of unknown length
* 35458 38012: contig of 2555 bp in length
* 38013 38112: gap of unknown length
* 38113 40338: contig of 2226 bp in length
* 40339 40438: gap of unknown length
* 40439 44051: contig of 3613 bp in length
* 44052 44151: gap of unknown length
* 44152 46294: contig of 2143 bp in length
* 46295 46394: gap of unknown length
* 46395 48845: contig of 2451 bp in length
* 48846 48945: gap of unknown length
* 48946 51753: contig of 2808 bp in length
* 51754 51853: gap of unknown length
* 51854 55915: contig of 4062 bp in length
* 55916 56015: gap of unknown length
* 56016 60468: contig of 4453 bp in length
* 60469 60568: gap of unknown length
* 60569 63640: contig of 3072 bp in length
* 63641 63740: gap of unknown length
* 63741 67888: contig of 4148 bp in length
* 67889 67988: gap of unknown length
* 67989 73014: contig of 5026 bp in length
* 73015 73114: gap of unknown length
* 73115 77313: contig of 4199 bp in length
* 77314 77413: gap of unknown length
* 77414 83786: contig of 6373 bp in length
* 83787 83886: gap of unknown length
* 83887 89590: contig of 5704 bp in length
* 89591 89690: gap of unknown length
* 89691 93846: contig of 4156 bp in length
* 93847 93946: gap of unknown length
* 93947 98126: contig of 4180 bp in length
* 98127 98226: gap of unknown length
* 98227 104258: contig of 6032 bp in length
* 104259 104358: gap of unknown length
* 104359 110994: contig of 6636 bp in length
* 110995 111094: gap of unknown length
* 111095 119838: contig of 8744 bp in length
* 119839 119938: gap of unknown length
* 119939 126339: contig of 6401 bp in length
* 126340 126439: gap of unknown length
* 126440 134888: contig of 8449 bp in length
* 134889 134988: gap of unknown length
* 142709: contig of 7721 bp in length

* 142710 142809: gap of unknown length
* 142810 149818: contig of 7009 bp in length
* 149819 149918: gap of unknown length
* 149919 158719: contig of 8801 bp in length
* 158720 158819: gap of unknown length
* 158820 169868: contig of 11049 bp in length
* 169869 169968: gap of unknown length
* 169969 180390: contig of 10422 bp in length
* 180391 180490: gap of unknown length
* 180491 191272: contig of 10782 bp in length
* 191273 191372: gap of unknown length
* 191373 209667: contig of 18295 bp in length
* 209668 209767: gap of unknown length
* 209768 226303: contig of 16536 bp in length.
FEATURES
 Location/Qualifiers
 source
 1. .226303
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /clone="CH230-7C10"
BASE COUNT 57875 a 53353 c 52917 g 57041 t 5117 others
Query Match 51.6%; Score 881.8; DB 2; Length 226303;
Best Local Similarity .98.4%; Pred. No. 2.2e-228;
Matches 922; Conservative 0; Mismatches 12; Indels 3; Gaps 3;
QY 775 CGCAAGGACACGAACTTTCTGCCCTTGTCACTCCTGCCTCTGCTGATGCTCTCTGTCCTCA 834
 |||||
Db 112250 CCCAGGACACGAACTTTCTGCCCTTGTCACTCCTGCCTCTGCTGATGCTCTCTGTCCTCA 112309
QY 835 ATTGGAGCACAGCCTGGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAG 894
 |||||
Db 112310 ATTGGAGCACAGCCTGGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAG 112369
QY 895 TGTCCAACCAGAACCGATTCTGCCAATCTGGAGATCCAAAGCCGCTGTGTCTGCCAGAC 954
 |||||
Db 112370 TGTCCAACCAGAACCGATTCTGCCAATCTGGAGATCCAAAGCCGCTGTGTCTGCCAGAC 112429
QY 955 CCTGCCCTGGCAGGACAGGACAGCTCATGGAACAGTGTCTTCTAAGGCCAACTGGGGA 1014
 |||||
Db 112430 CCTGCCCTGGCAGGACAGGACAGCTCATGGAACAGTGTCTTCTAAGGCCAACTGGGGA 112489
QY 1015 TGGCGATACAGGGCCTGCCATCCTCAGCAAAATGACCCCTAGGACCAAGCCCTGGACTGCTG 1074
 |||||
Db 112490 TGGCGATACAGGGCCTGCCATCCTCAGCAAAATGACCCCTAGGACCAAGCCCTGGACTGCTG 112549
QY 1075 GTAGATGCTCTTCTCCATGCTCTTGGTGCAGTTAACTGTCTCCTGCTGGATTCACTGTGT 1134
 |||||
Db 112550 GTAGATGCTCTTCTCCATGCTCTTGGTGCAGTTAACTGTCTCCTGCTGGATTCACTGTGT 112609
QY 1135 AGAGCCACTGAGCGATCCCTGCTCTGTCTGAGGTA-GGCGGAGAGGAGTACCAGCTCCAG 1193
 |||||
Db 112610 AGAGCCACTGAGCGATCCCTGCTCTGTCTGAGGTAAGGGCGGAGAGGTGACCAGCTCCAG 112669
QY 1194 TTCTCTGGTTCAG CCTGGAATTCGGGTTCCTCCTGGCTCATTCCTCAAAACATCCCTGT 1252
 |||||
Db 112670 GTCTCTGGTTCAG CCTGGGATTCGGGTTCCTCCTGGCTTCCTCTCAAAACATCCCTGT 112729
QY 1253 AAAAAAGGACA-ACCAAAAAGACCTTTAAACCTAGGCTATAGGCGCAAAACCTGGCCAC 1311
 |||||
Db 112730 AAAAAAGGTCACACCAAAAAGACCTTTAAACCTAGGCTATAGGCGCAAAACCTGGCCAC 112789
QY 1312 CGTGTGGGATAAGGTCAATGTTAGGACCAGACAGCAGATTGCTGAAACTTCCAATTC 1371
 |||||
Db 112790 CGTGTGGGATAAGGTCAATGTTAGGACCAGACAGCAGATTGCTGAAACTTCCAATTC 112849
QY 1372 CCTTCTTGGACTTCTGTATGCTTGTCCCAAAAGATGATGAATGAATCGTAAGTGTACCT 1431
 |||||
Db 112850 CCTTATTGGACTTCTGTATGCTTGTCCCAAAAGATGATGAATGAATCGTAAGTGTACCT 112909
QY 1432 TCCCTGACCTGAGAACACCCCTGCCTCGGGGAAGTATTTCAGGGGCGAGAATTCTCTGTGA 1491
 |||||
Db 112910 TCCCTGACCTGAGAACACCCCTGCCTGCTCGGGGAAGTATTTCAGGGGCGAGAATTCTCTGTGA 112969

QY 1492 ACATGAAGAGATGAATCACACTGTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGACCT 1551
|||||
Db 112970 ACATGAAGAGATGAATCACACTGTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGACCT 113029

QY 1552 TTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTC 1611
|||||
Db 113030 TTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTC 113089

QY 1612 TGGCAGGCCAGGGCCCTTTCTCTTCAGCATGAGAAAGACAGAGGACAGAGTACTCTCC 1671
|||||
Db 113090 TGGTAGGCCAGGGCCCTTTCTCTTCAGCATGAGAAAGACAGAGGACAGAGTACTCTCC 113149

QY 1672 TCTGAGGACTAGTCTAGCCTAGAATAAACACCCAAA 1708
|||||
Db 113150 TCTGAGGACTAGTCCAGCCTAGATAAACAACCCAAA 113186

RESULT 8
AL731698
LOCUS AL731698.10 GI:21214309 61072 bp DNA linear ROD 24-MAY-2002
DEFINITION Mouse DNA sequence from clone RP23-161B3 on chromosome 2, complete
sequence.
ACCESSION AL731698
VERSION AL731698.10
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 61072)
AUTHORS Wallis,J.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk

COMMENT
On May 25, 2002 this sequence version replaced gi:21213601.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30);
an attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-161B3 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6.

FEATURES
source
1..61072
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-161B3"
/clone_lib="RPCI-23"
BASE COUNT 15405 a 15368 c 15033 g 15266 t
ORIGIN

Query Match 37.8%; Score 645; DB 10; Length 61072;
Best Local Similarity 85.6%; Pred. No. 5.3e-164;
Matches 804; Conservative 0; Mismatches 115; Indels 20; Gaps 7;
QY 775 CGCAAGGACACCAACTTTCTGCCCTTGTCACCTCCTCGCTGATGCTCTTGTCACAA 834

Db 16998 CCCAGGACACCAACTTTCTGCCCTTGTCACCTCCTGCATCTGCCGATGGCCCTGTCCAA 17057
|||||
QY 835 ATTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGGGGCTGGGCATAGCCACCCGAG 894
|||||
Db 17058 ACTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGGGGTTGGGCATAGCCACCCGAG 17117

QY 895 TGTCCAACCAAGACCGATTCTTGCCAACTGGAGATCCACGCCCTGTGTCTGCCAGAC 954
|||||
Db 17118 TATCCAACCAAGACCGATTCTTGCCAACTGGAGATCCAGGCTGCCTGTGTCTGCCAGAC 17177

QY 955 CCTGCCTGGCAGCCAGGAGCACAGCTCATGGAACAGTGTCTTCTAAGGCCA-ACTGGGG 1013
|||||
Db 17178 CCTGCCTGGCATCCAGGAGCACGGCTCATGGAACAGTGTCTTAGAGCCATTGGCGGG 17237

QY 1014 ATCGGGATACAGGGCCTGCCATCCTCAGCAAATGACCCCTAGGACCAGGCCCTGGACTGCT 1073
|||||
Db 17238 ATGTGGATACAGGGCCTGCCATTCTCAGCAAATGTCCCTAGGACCAGGCCCTGGACTGAT 17297

QY 1074 GGTAGATGCTCTTCTCCATGCTTCTGGCTGCAGTTAACTGTCTCTGGATTCACTGTG 1133
|||||
Db 17298 GGTAGATGCCCCCTCTCCATGCTTCTGGCTGCAGTTAACTGTCTGGTGGATTCACTGTG 17357

QY 1134 TAGAGCCACTGAGCGATCCCTGCTCTGTCTGAGGTAGGGGAGCAGGTGACCAGTCCAG 1193
|||||
Db 17358 CAGAGCCTCTGAGCGATCCCTGCTCTGTCTGAGGTGGGGAAGCAGGTGACCAGTCCAT 17417

QY 1194 TTCTCTGG--TTCAGCCTGGAATTCCTGGGTTCTCCTGGCTCATTCCTCAAAACATCCCTG 1251
|||||
Db 17418 TTCTCTGGATTCTGACCCAGGCTTCTGGGTTCTCCTGCTAGTTCCTCAAAACTTCCCTG 17477

QY 1252 TACAAAAAGGACAAACCAAAAGACCTTTAAACCTAGGCTATACTGGGCAACCTGGCCAC 1311
|||||
Db 17478 TATGAAAAGGACAAACCAAAAGGACCTTTAAAGCTAAGCTGTACTGGGCAAGCCTGGCCAC 17537

QY 1312 CGTGTGGGGATAAGGTCAATGTTAGG-ACCAGACAGCAGATTGCCTGAAACTTCCAAAT 1370
|||||
Db 17538 CATGTGGGGATAGTGACAGTAATAGGTACCAGGCAGCAGATTGCCTGAAACATCCAGGT 17597

QY 1371 CCCTTCTTGGACTTCTGTATGCTTGTCCTCCCAAAGATGATGAATGAACCTCGTAAGTGACC 1430
|||||
Db 17598 CCCTTCTTGGACTTCTATGCTTGT-CCCAAAGATTATGGGTGACCTTGTAAGTGTCGC 17656

QY 1431 TTCCTGACCTGAGAACACCCCTGCCTGCTCGGGAAAGTATTCAGGGGCAGAGAATTCCTGTG 1490
|||||
Db 17657 TTCTCTGATCTGAGAACACCCCTGCCGGCTGGGA-----AGAATTTCTGGG 17703

QY 1491 AACATGAAGAGAT-GAATCACACTGTCTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAG 1549
|||||
Db 17704 AACATGAAGAGATGGAATCACACTATTCTTAAGAGCGTTTGCCAAGTCCAGGAACCTTGAC 17763

QY 1550 CTTGTATTTTCAGGAATGCACATCTCTTAAGCACCTCGCAAAACAGGAAGGCTCCACACC 1609
|||||
Db 17764 CTTGTATTTGTAAAAATACACATCTCTTAAATGCTCACAAGCAAG-AGGCTCCACACT 17822

QY 1610 TCTGGCAGGCCAGGGCCTTTCTTTCAGCATGAGAAAGACAAAGGACAGAGTACTCT 1669
|||||
Db 17823 TCTGGCAGGCCAGGGCCTTTCTTTCAGCATGAGAGAGACAAAGAACAGTAGATACCCT 17882

QY 1670 CCTCTGGAGGACTAGTCTAGCTAGATAAATAAACACCCAAA 1708
|||||
Db 17883 CCTCTGGAGGACTGGCCCCGCTGTGGAATAAACAACCCAAA 17921

RESULT 9
AL669906
LOCUS AL669906 216757 bp DNA linear HTG 24-JUL-2002
DEFINITION Mus musculus chromosome 2 clone RP23-217C2, *** SEQUENCING IN
PROGRESS ***, 21 unordered pieces.
ACCESSION AL669906
VERSION AL669906.5 GI:21955520
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Mus musculus.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 216757)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 25, 2002 this sequence version replaced gi:18181793.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BM217C2
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 211865 bases at least Q40
Consensus quality: 213403 bases at least Q30
Consensus quality: 214139 bases at least Q20
Insert size: 214757; sum-of-contigs
Insert size: 234243; 1.8% error; agarose-fp
Quality coverage: 6.29x in Q20 bases; sum-of-contigs Quality
coverage: 5.90x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 4000: contig of 4000 bp in length
* 4001 4100: gap of 100 bp
* 4101 15267: contig of 11167 bp in length
* 15268 15367: gap of 100 bp
* 15368 34970: contig of 19603 bp in length
* 34971 35070: gap of 100 bp
* 35071 39671: contig of 4601 bp in length
* 39672 39771: gap of 100 bp
* 39772 46360: contig of 6589 bp in length
* 46361 46460: gap of 100 bp
* 46461 49809: contig of 3349 bp in length
* 49810 49909: gap of 100 bp
* 49910 54509: contig of 4600 bp in length
* 54510 54609: gap of 100 bp
* 54610 65989: contig of 11380 bp in length
* 65990 66089: gap of 100 bp
* 66090 84635: contig of 18546 bp in length
* 84636 84735: gap of 100 bp
* 84736 87399: contig of 2664 bp in length
* 87400 87499: gap of 100 bp
* 87500 98601: contig of 11102 bp in length
* 98602 98701: gap of 100 bp
* 98702 105522: contig of 6821 bp in length
* 105523 105622: gap of 100 bp
* 105623 119773: contig of 14151 bp in length
* 119774 119873: gap of 100 bp
* 119874 134552: contig of 14679 bp in length
* 134553 134652: gap of 100 bp
* 134653 141202: contig of 6550 bp in length
* 141203 141302: gap of 100 bp
* 141303 145808: contig of 4506 bp in length
* 145809 145908: gap of 100 bp
* 145909 151572: contig of 5664 bp in length
* 151573 151672: gap of 100 bp
* 151673 169686: contig of 18014 bp in length
* 169687 169786: gap of 100 bp
* 169787 200095: contig of 30309 bp in length

FEATURES

source
* 200096 200195: gap of 100 bp
* 200196 213877: contig of 13682 bp in length
* 213878 213977: gap of 100 bp
* 213978 216757: contig of 2780 bp in length.
Location/Qualifiers
1..216757
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-217C2"
/clone_lib="RPCI-23"
1..4000
/note="assembly_fragment:03429
clone_end:SP6
vector_side:left"
4101..15267
/note="assembly_fragment:00814
fragment_chain:1"
15368..34970
/note="assembly_fragment:00812
fragment_chain:1"
35071..39671
/note="assembly_fragment:03544
fragment_chain:1"
39772..46360
/note="assembly_fragment:01438
fragment_chain:1"
46461..49809
/note="assembly_fragment:03410
fragment_chain:2"
49910..54509
/note="assembly_fragment:00811
fragment_chain:2"
54610..65989
/note="assembly_fragment:02217
fragment_chain:2"
66090..84635
/note="assembly_fragment:03529
fragment_chain:2"
84736..87399
/note="assembly_fragment:01019
fragment_chain:3"
87500..98601
/note="assembly_fragment:01170
fragment_chain:3"
98702..105522
/note="assembly_fragment:02829
fragment_chain:3"
105623..119773
/note="assembly_fragment:01476
fragment_chain:4"
119874..134552
/note="assembly_fragment:00505
fragment_chain:4"
134653..141202
/note="assembly_fragment:01951
fragment_chain:4"
141303..145808
/note="assembly_fragment:02637
fragment_chain:5"
145909..151572
/note="assembly_fragment:02352
fragment_chain:5"
151673..169686
/note="assembly_fragment:02207
fragment_chain:5"
169787..200095
/note="assembly_fragment:02849
fragment_chain:6"
200196..213877
/note="assembly_fragment:01052
fragment_chain:6"
213978..216757


```

the fourth carboxy-terminal domain present in other
members of the CCN family"
/codon_start=1
/product="connective tissue growth factor-like protein
precursor"
/protein_id="AAC70350.1"
/db_xref="GI:3462836"
/translation="MRGTPKTHLLAFSLCLLSKVRTQLCPTCTCPWPPPRCPPLGVP
LVLDGCGCRRVCARRLGEPCDQLHVCDAQGLVCQPGAGGGRGALCLLAEDDSCEV
NGRLYREGETFQPHCSIRCRCEGDGGFTCVPLCSEDVRLPSWDCHPRRVEVLGKCCPE
WVCGGGGLGTQPLPAQGPQFSLGVSSLPVGPVCPPEWSTAWGPCSTTCGLGMATRVSN
QNRECRLETQRRLLCLSRPCPPSRGRSPQNSAF"
235 a 418 c 389 g 241 t
BASE COUNT
ORIGIN

```

		Query Match	32.9%;	Score 561.4;	DB 9;	Length 1283;
		Best Local Similarity	78.0%;	Pred. NO. 2.2e-141;		
		Matches 701;	Conservative 0;	Mismatches 196;	Indels 2;	Gaps 2;
QY	243	GGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCCTTCCCTCTGCTCTCTC 302				
DB	3	GGGGACATGAGAGGCACACCGAAGACCCCACTCTCTGGCCCTTCTCCCTCTCTGCTCTCTC 62				
QY	303	TCAATGGTGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTTGGACACCAACCCAG 362				
DB	63	TCAAAGGTGGGTACCCAGCTGTGCCCGACACCACTGTACCTGTCCCTTGGCCACTTCCCGCA 122				
QY	363	TGCCCCACAGGGGGTACCCCTGGTGGCTGGATGGCTGTGGCTGTCTGTAAAGTGTGTGCACGG 422				
DB	123	TGCCCGCTGGGAGTACCCCTGGTGGTGGATGGCTGTGGCTGTCTGCGGGGTATGTGCACGG 182				
QY	423	AGGTGGGGAGTCTCGACCACTGCTGCTGCGACCCAGCCAGCCAGGGCCCTGGTTGT 482				
DB	183	CGGCTGGGGAGCCCTGCGACCAACTCCACCTCTGCGACGCCAGCCAGGGCCCTGGTCTGC 242				
QY	483	CAGCCTGGSGAGGCCCTGGCGGCCATGSGSGTGTGTGTCTCTTGGATGAGGATGACGGT 542				
DB	243	CAGCCCGGGCAGGACCCGGTGGCCGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGC 302				
QY	543	AGCTGTAGGTGAATGGCCGAGGTACCTGATGGAGAGACCTTTAAACCAATTGCAGG 602				
DB	303	AGCTGTAGGTGAACGGCCGCCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 362				
QY	603	GTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCGCTGTGCAGTGAGGATGTG 662				
DB	363	ATCCGCTGCCGCTGCGAGGACGGCGGCTTACCTGCGTCCGCTGTGCAGCAGGATGTG 422				
QY	663	CGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGAATACAGGTGCCAGGAAAGTGTCTGC 722				
DB	423	CGGCTGCCAGCTGGGACTGCCCCCAACCCAGGAGGGTTCGAGGTCTTGGGCAAGTGTCTGC 482				
QY	723	CCCGAGTGGGTATGTGACCAGGGAGTGACACCGGGCGATCCAGCGCTCCACGCGCCCAAGGA 782				
DB	483	CCTGAGTGGTGTGCGGCCAAGGAGGGGGACTGGGGACCCAGCCCCCTTCCAGCCCCAAGGA 542				
QY	783	CACCAACTTCTGCCCCCTGTCTACTCTGCTGCTGATGCTCCTTGTCTCAAAATTGGAGC 842				
DB	543	CCCCAGTTTCTGGCCCTTGTCTCTTCCCTGCCCTGGTGTCTCCCTGCCCAAGATGGAGC 602				
QY	843	ACAGCCTGGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAC 902				
DB	603	ACGGCCTGGGGACCCCTGCTCGACCACCTGTGGGCTGGGCATGGCCACCCGGGTGTCCAAC 662				
QY	903	CAGAACCGATTCTGCCAACTGGAGATCCAAACCGCCCTGTGTCTGCCACAGACCCCTGCCTG 962				
DB	663	CAGAACCGCTTCTGCCGACTGGAGACCCAGCGCCGCTGTGCCTGTCCAGGGCCCTGCCCA 722				
QY	963	GCAGCCAGGAGCCACAGCTCATGGAAACAGTGTCTTCTA-AGGCCAACTGGGATGCGGAT 1021				
DB	723	CCCTCCAGGGTTCGAGTCCACAAACAGTGCCTTCTAGAGCCGGGTGGGAATGGGGAC 782				
QY	1022	ACAGGGCCTGCCATCCTCAGCAAAATGACCCCTAGGACCAAGGCCCTGGACTGCTGGTAGATG 1081				

[illegible]

/db_xref="GI:3328192"
/translation="MRGTPKTHLLAFSLCLLSKVRTQLCPPTCPWPPRCPLGVP
LVLDGCGCCRVCAARLGPEDQLHVCDSQGLVCQGAGPGRGALCLLAEDDSSCEV
NGRLYREGETFQPHCSIRCECDQGFCTCVPLCSEDVRLPSWDCPHPRRVEVLGKCCPE
WVCGGGGLGTQPLPAQGPQFSGLVSSLPVPGVPCPEWSTAWGPCSTTCGLGMATRVSN
QNRFCRLFTQRRILCLSRPCPPSRGRSPQNSAF"

Search completed: July 29, 2003, 00:48:27
Job time : 4752.96 secs

BASE COUNT	261 a	418 c	387 g	242 t	1 others
ORIGIN					
Query Match		32.9%	Score 561.4;	DB 9;	Length 1309;
Best Local Similarity		78.0%;	Pred. No. 2.2e-141;		
Matches 701;	Conservative	0;	Mismatches 196;	Indels 2;	Gaps 2;
QY	243	GGTGACATGAGGGGAGCCCACTGATCCATCTTCTTGGCCACTTCCTTCCTCTGCCTTCTC	302		
Db	1	GGGACATGAGAGCACACCGAAGACCCACCTCTGGCCTTCTCCCTCCTCTGCCTCCTC	60		
QY	303	TCAATGGTGTGTCGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAG	362		
Db	61	TCAAAGGTGCGTACCCAGCTGTGCCCGACACATGTACCTGCCCTGGCCACCTCCCCGA	120		
QY	363	TGCCCACAGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTATAAGTGTGTGCACGG	422		
Db	121	TGCCCGCTGGGAGTACCCCTGGTGGTGGATGGCTGTGGCTGTGCCGSTATGTGCACGG	180		
QY	423	AGGCTGGGGAGTCCTGCGACCACTGCATGTCTGCGACCCCAAGCCAGCCAGGCCCTGTTGT	482		
Db	181	CGGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGC	240		
QY	483	CAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT	542		
Db	241	CAGCCCGGGGAGGACCCGGTGGCCGGGGGCCCTGTGCCTTGGCAGAGGACGACAGC	300		
QY	543	AGCTGTGAGGTGAATGGCCGCGAGGTACCTGGATGGAGAGACCTTTAAACCCAATTGCAGG	602		
Db	301	AGCTGTGAGGTGAACGGCCGCCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGC	360		
QY	603	GTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGCAGTGAGGATGTG	662		
Db	361	ATCCGCTGCCGCTGCGAGGACGGGGCTTCACCTGCGTGCCTGTGCAGCGAGGATGTG	420		
QY	663	CGGCTGCCCAGCTGGGACTGCCCCACGCCCAAGAGAATACAGGTGCCAGGAAAGTGTGC	722		
Db	421	CGGCTGCCCAGCTGGGACTGCCCCACCCCAAGAGGGTTCGAGGTCTCTGGGCAAGTGTGC	480		
QY	723	CCCAGTGGGTATGTGACCAGGGAGTGACACCGGGCGATCCAGCGTCCACGGCGCAAGGA	782		
Db	481	CCTGAGTGGGTGCGGCCCAAGAGGGGGACTGGGGACCCAGCCCTTCCAGCCCCAAGGA	540		
QY	783	CACCAACTTCTGCCCTTGTGCTACTCCTGCTGTGCTGATGCTCCTTGTCCAAATTGGAGC	842		
Db	541	CCCCAGTTTCTGGCCCTTGTCTCTTCCCTGGCCCTTGGTGTCCCTGCCAGAAATGGAGC	600		
QY	843	ACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGCATAGCCACCGAGTGTCCAAC	902		
Db	601	ACGGCCTGGGGACCCCTGCTCGACCACTGTGGCTGGCATGGCCACCCGGGTGTCCAAC	660		
QY	903	CAGAACCGATTCTGCCAACTGGAGATCCAAACGCCGCCCTGTGTCTGCCAGACCCCTGCCTG	962		
Db	661	CAGAACCGCTTCTGCCGACTGGAGACCCAGCGCCGCCCTGTGCCTGTCCAGGCCCTGCCCA	720		
QY	963	GCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTA-AGGCCAACTGGGGATGCGGAT	1021		
Db	721	CCCTCCAGGGGTGCGAGTCCACAAAACAGTGCCTTCTAGAGCCGGGCTGGGAATGGGGAC	780		
QY	1022	ACAGGGCCTGCCATCCTCAGCAAATGACCTAGGACCAAGGCCCTGGACTGCTGGTAGATG	1081		
Db	781	ACGGTGTCCACCATCCCACTGGTGGCGCTGTGCCTGGGCCCTGGGCTGATGGAAGATG	840		
QY	1082	CT-CTTCTCCATGCTTGGCTGCAGTTAACTGTCTCTGCTTGGATTCACTGTGTAGAGC	1139		
Db	841	GTCCGTGCCCCAGGCCCTTGGCTGCAGGCACACACTTAGCTTGGGTCCACCATGCAGAAC	899		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2003, 09:03:08 ; Search time 452.287 Seconds
(without alignments)
8504.365 Million cell updates/sec

Title: US-10-010-408-1
Perfect score: 1708
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GCCTAGATAAACACCCAAA 1708

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1708	100.0	1708	20	AAZ07516 Rat HICP polypepti
2	1278	74.8	1734	20	AAZ07516 Mouse WISP-2 prote
3	753	44.1	753	20	AAZ07517 Rat HICP polypepti
4	681	39.9	681	20	AAZ07521 Rat HICP mature po
5	657	38.5	753	20	AAZ07521 Mouse WISP-2 prote
6	583.4	34.2	1522	20	AAZ07521 Human growth facto
7	566.8	33.2	1337	22	AAH46952 Human secreted pro
8	566.6	33.2	1352	22	AAH46936 Human secreted pro
9	561.4	32.9	1266	21	AAA30048 Human PRO261 nucle

10	561.4	32.9	1266	22	AAS21403 Human cDNA sequenc
11	561.4	32.9	1266	22	AAF60368 PRO261 coding sequ
12	561.4	32.9	1266	22	AAC97451 Human angiogenesis
13	561.4	32.9	1285	19	AAV29260 Human connective t
14	561.4	32.9	1293	20	AAZ07516 Human WISP-2 prote
15	561.4	32.9	1309	22	AAH28214 Connective tissue
16	557	32.6	1257	20	AAZ07516 EGF-like homologue
17	550.4	32.2	1267	21	AAZ07516 Human PRO261 cDNA
18	528.8	31.0	841	20	AAZ07516 Human WISP-2 prote
19	508.4	29.8	750	20	AAZ07516 Human WISP-2 prote
20	501.2	29.3	738	20	AAZ07516 Human WISP-2 prote
21	408.2	23.9	439	24	ABL59575 Rat OST23 gene fra
22	210	12.3	210	20	AAZ07519 Rat HICP IGFBP dom
23	203.4	11.9	2136	22	AAZ07519 Human full-length
24	195.8	11.5	13255	22	AAZ07519 Human immune/haema
25	177	10.4	177	20	AAZ07518 Rat HICP VWC doma
26	174	10.2	174	20	AAZ07520 Rat HICP TSPI doma
27	169.8	9.9	586	22	ABA59703 Human foetal liver
28	169.8	9.9	586	22	ABA28232 Probe #6698 for ge
29	169.8	9.9	586	22	AAK07972 Human brain expres
30	169.8	9.9	586	22	AAK33842 Human bone marrow
31	169.8	9.9	586	22	AAI16552 Probe #6485 for ge
32	169.8	9.9	586	22	AAI39568 Probe #8254 used t
33	169.8	9.9	586	24	ABS08689 Human genome-deriv
34	163.8	9.6	2075	16	AAT04226 Connective tissue
35	163.8	9.6	2075	18	AAT51234 Connective tissue
36	163.8	9.6	2075	18	AAZ07516 Human connective t
37	163.8	9.6	2075	19	AAV38085 Human connective t
38	163.8	9.6	2075	20	AAZ07516 Human connective t
39	163.8	9.6	2075	21	AAZ07516 Human connective t
40	163.8	9.6	2075	21	AAZ07516 Human connective t
41	163.8	9.6	2075	22	AAZ07516 Human connective t
42	163.8	9.6	2075	22	AAZ07516 Human connective t
43	163.8	9.6	2075	22	AAZ07516 Human connective t
44	163.8	9.6	2075	24	ABK64553 Human benign prosta
45	163.8	9.6	2312	22	AAH28213 Nucleotide sequenc

ALIGNMENTS

RESULT 1	
AAZ07516	
ID	AAZ07516 standard; cDNA; 1708 BP.
XX	
AC	AAZ07516;
XX	
DT	26-NOV-1999 (first entry)
XX	
DE	Rat HICP polypeptide encoding cDNA.
XX	
KW	Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW	cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX	
OS	Rattus sp.
XX	
PN	WO9947556-A2.
XX	
PD	23-SEP-1999.
XX	
PF	18-MAR-1999; 99WO-US05999.
XX	
PR	19-MAR-1998; 98US-0044273.
XX	
PA	(TUFT) TUFTS COLLEGE.
XX	
PI	Castellot JJ;
XX	
DR	WPI; 1999-562060/47.
DR	P-PSDB; AAY27434.
XX	
PT	Nucleic acid sequences encoding rat heparin-induced CCN-like protein, used in methods to identify modulators or in diagnostic applications

XX Claim 2; Fig 1; 108pp; English.

XX This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein.

CC Agents that stimulate or inhibit HICP protein activity or expression,

CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to

CC modulate cell-associated activity. HICP modulators can be used to treat

CC disorders characterized by aberrant HICP protein activity or expression.

CC Probes capable of hybridizing to HICP mRNA or antibodies specific for

CC HICP can be used to detect HICP activity in a biological sample. HICP

CC can be used to treat disorders, such as a cardiovascular or fibrotic

CC disorder, characterized by aberrant cell proliferation.

XX

SQ Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Query Match 100.0%; Score 1708; DB 20; Length 1708;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGCTTCTGATCTCCAGAGGACCCTGGGTTGGACAGGGCCCTTGGCAAGGCTGCAGCC 60

DB 1 GACGCTTCTGATCTCCAGAGGACCCTGGGTTGGACAGGGCCCTTGGCAAGGCTGCAGCC 60

QY 61 GCTGGGCAGTGGCTTGAATGGAGGTCTTTTATTACTGGGAAGCTGAGGAGCTTAAGAGGCTC 120

DB 61 GCTGGGCAGTGGCTTGAATGGAGGTCTTTTATTACTGGGAAGCTGAGGAGCTTAAGAGGCTC 120

QY 121 CTGTCAGCTTGTCTAAAGTCTTAGCACCTTGTGGTGGCTTGGCTTCACACACTGTCAGA 180

DB 121 CTGTCAGCTTGTCTAAAGTCTTAGCACCTTGTGGTGGCTTGGCTTCACACACTGTCAGA 180

QY 181 CACCTTCGTGGTGGCTCCACGGCCCTCACCTTCAAGCTTGAAGCTGGCTCCACAAGGGAC 240

DB 181 CACCTTCGTGGTGGCTCCACGGCCCTCACCTTCAAGCTTGAAGCTGGCTCCACAAGGGAC 240

QY 241 ACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCCTTCCCTTGCCTTC 300

DB 241 ACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCCTTCCCTTGCCTTC 300

QY 301 TCTCAATGGTGTGTCOCAGCTGTGCCGGACACCCCTGTACCTGTCCCTGGACACCAACCCC 360

DB 301 TCTCAATGGTGTGTCOCAGCTGTGCCGGACACCCCTGTACCTGTCCCTGGACACCAACCCC 360

QY 361 AGTGCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTAAAGTGTGTGCAC 420

DB 361 AGTGCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTAAAGTGTGTGCAC 420

QY 421 GGAGGCTGGGGAGTCTTCCGACCACTGTCATGTCTGCGACCCCAAGCCGCTGGTTT 480

DB 421 GGAGGCTGGGGAGTCTTCCGACCACTGTCATGTCTGCGACCCCAAGCCGCTGGTTT 480

QY 481 GTCAGCCTGGGGCAGGCCCTGGGGCCATGGGCTGTGTCTCTTTGGATGAGGATGACG 540

DB 481 GTCAGCCTGGGGCAGGCCCTGGGGCCATGGGCTGTGTCTCTTTGGATGAGGATGACG 540

QY 541 GTAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCA 600

DB 541 GTAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCA 600

QY 601 GGGTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCGCTGTGCAGTGGATG 660

DB 601 GGGTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCGCTGTGCAGTGGATG 660

QY 661 TCGGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGATAACAGGTGCCAGGAAGTGCT 720

DB 661 TCGGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGATAACAGGTGCCAGGAAGTGCT 720

QY 721 GCCCGAGTGGGTATGTGACCAAGGAGTGACACGGCGGATCCAGCGCTCCACGGCGCAAG 780

DB 721 GCCCGAGTGGGTATGTGACCAAGGAGTGACACGGCGGATCCAGCGCTCCACGGCGCAAG 780

QY 781 GACACCAACTTCTGCCCCTTGTCACTCCCTGCTGCTGATGCTCCTTGTCCAAATTTGGA 840

DB 781 GACACCAACTTCTGCCCCTTGTCACTCCCTGCTGCTGATGCTCCTTGTCCAAATTTGGA 840

Db 781 GACACCAACTTCTGCCCCTTGTCACTCCCTGCTGCTGATGCTCCTTGTCCAAATTTGGA 840

QY 841 GCACAGCCTGGGCCCCCTGCTCAACCACCTGTGTGGGCTGGGCATAGCCACCCGAGTGTCCA 900

Db 841 GCACAGCCTGGGCCCCCTGCTCAACCACCTGTGTGGGCTGGGCATAGCCACCCGAGTGTCCA 900

QY 901 ACCAGAACCGGCTTCTGCCAACTGGAGATCCAACGCCCGCTGTGTCTGCCAGACCCCTGCC 960

Db 901 ACCAGAACCGGCTTCTGCCAACTGGAGATCCAACGCCCGCTGTGTCTGCCAGACCCCTGCC 960

QY 961 TGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCCAACTGGGATGCCGA 1020

Db 961 TGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCCAACTGGGATGCCGA 1020

QY 1021 TACAGGGCCCTGCCATCCTCAGCAAAATGACCCCTAGGACCCCTGGACTGCTGGTAGAT 1080

Db 1021 TACAGGGCCCTGCCATCCTCAGCAAAATGACCCCTAGGACCCCTGGACTGCTGGTAGAT 1080

QY 1081 GCTCTTCTCCATGCTCTTGGCTGCAGTTAACTGCTTGGATTCACTGTGTAGAGCC 1140

Db 1081 GCTCTTCTCCATGCTCTTGGCTGCAGTTAACTGCTTGGATTCACTGTGTAGAGCC 1140

QY 1141 ACTGAGCGATCCTGCTCTGTGTAGGTAGGGGAGCAGGTGACCAAGTTCAGTGTGTAGAGCC 1200

Db 1141 ACTGAGCGATCCTGCTCTGTGTAGGTAGGGGAGCAGGTGACCAAGTTCAGTGTGTAGAGCC 1200

QY 1201 GTTCAGCCTGGAACTTCTGGGTTCTCCTGGCTCAATCCTCAAAACATCCCTGTACAAAAG 1260

Db 1201 GTTCAGCCTGGAACTTCTGGGTTCTCCTGGCTCAATCCTCAAAACATCCCTGTACAAAAG 1260

QY 1261 GACAACCAAAAAGACCTTTAAACCTAGGCTATCTAGGCAAAACCTGGCCACCGTGTGGG 1320

Db 1261 GACAACCAAAAAGACCTTTAAACCTAGGCTATCTAGGCAAAACCTGGCCACCGTGTGGG 1320

QY 1321 GATAAGGTCAATGTAGGACCAGACAGCAGATTCCTGAAACTTCCAATTCCTTCTTGG 1380

Db 1321 GATAAGGTCAATGTAGGACCAGACAGCAGATTCCTGAAACTTCCAATTCCTTCTTGG 1380

QY 1381 ACTTCTGTATGCTGTGTCCTCCCAAGATGATGAATCGTAAGTGTACCTTCCCTGACC 1440

Db 1381 ACTTCTGTATGCTGTGTCCTCCCAAGATGATGAATCGTAAGTGTACCTTCCCTGACC 1440

QY 1441 TGAGAACACCCCTGCTGCTCGGGAAGTATTCAGGGGCGAGAATTCCTGTGAACATGAAGA 1500

Db 1441 TGAGAACACCCCTGCTGCTCGGGAAGTATTCAGGGGCGAGAATTCCTGTGAACATGAAGA 1500

QY 1501 GATGAATCACACTGCTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTGTATTTT 1560

Db 1501 GATGAATCACACTGCTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTGTATTTT 1560

QY 1561 CAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTTGGCAGGCC 1620

Db 1561 CAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTTGGCAGGCC 1620

QY 1621 AGGGCCTTTCTCTTCAGCATGAGAAAGACAGGGACAGCAGAGTACTCTCTCTGGAGGA 1680

Db 1621 AGGGCCTTTCTCTTCAGCATGAGAAAGACAGGGACAGCAGAGTACTCTCTCTGGAGGA 1680

QY 1681 CTAGTCTAGCCTAGAATAAACACCCCAA 1708

Db 1681 CTAGTCTAGCCTAGAATAAACACCCCAA 1708

RESULT 2

AAAX76488

ID AAAX76488 standard; DNA; 1734 BP.

XX

AC AAAX76488;

XX

DT 06-AUG-1999 (first entry)

XX Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.

XX

KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX Mus sp.
OS
XX
PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH) GENENTECH INC.
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI; 1999-337420/28.
DR P-PSDB; AAY17651.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Example 2; Page 178-179; 284pp; English.
XX
CC The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;

Query Match. 74.8%; Score 1278; DB 20; Length 1734;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 1520; Conservative 0; Mismatches 165; Indels 33; Gaps 11;

QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCG 62
DB 13 CGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCCCTTGGCAGGCTGCAGCTGC 72

QY 63 TG-GGCAGTGGCTTGAATGGAGGCTCTTTATTACTGGGAAGTGGAGGCTAAGAGGCTCC 121
DB 73 TGTGGCAGTAGCTTGGGATGGAGGCTCTTCTTGCTGGGAAGTGGAGGCTGAGAGGCTCC 132

QY 122 TGTCAG---CTGTCTCTAAAGTCTTAGCACTGTGGTGGCTTGGGCTTACACACTGTCA 178
DB 133 TGTACGGCTCCTGTCTCTAAACTCTTGGCACTTGGGCTGGGCTTGGGCTTACACACTGTCA 192

QY 179 GACACCTTCGTGGTGGCCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGG 238
DB 179 GACACCTTCGTGGTGGCCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGG 238

Db 193 GACACCTTCTTGGTGGCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGG 246
QY 239 ACACGGTGACATGAGGGGACGCCACATGATCCATCTTCTGGCCACTTCTCTCTGCT 298
Db 247 ACACGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCATTTCTCTCTG 306
QY 299 TCTCTCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACC 358
Db 307 TCTCTCAATGGTGTATTCCCAGCTGTGCCAGCACCCCTGTGCCTGTCTTGGACACC 366
QY 359 CCAGTGCCACAGGGGGTACCCCTGGTGGTGGATGGTGTGGCTGTGTAAGTGTGTGC 418
Db 367 CCAGTGCCACCGGGGTACCCCTGGTGGTGGATGGTGTGGCTGTGCGAGTGTGTGC 426
QY 419 ACGGAGGCTGGGGGAGTCTCTGCGACCACCTGCATGTCTGCGACCCACCCAGGGCCTGGT 478
Db 427 ACGGAGGCTGGGGGAGTCTCTGCGACCACCTGCATGTCTGCGACCCACCCAGGGCCTGGT 486
QY 479 TTGTACGCTTGGGGCAGGCCCTGGCGGCATGGGGCTGTGTCTCTTGGATGAGGATGA 538
Db 487 TTGTACGCTTGGGGCAGGCCCCAGTGGCCGTGGTGTGTGTGCTTCTCGAAGAGGATGA 546
QY 539 CGTAGCTGTGAGGTGAATGGCCGACGTGACCTGGATGGAGAGACCTTTAAACCCAATTG 598
Db 547 CGGGAGCTGTGAGGTGAATGGCCGACGTGACCTGGATGGGGAGACCTTTAAACCCAATTG 606
QY 599 CAGGGTCTCTGCCGCTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTCAGTGAGGA 658
Db 607 CAGGGTCTTGTGCCGCTGTGATGACGGTGGTGTACCTGCCTGCCGCTGTCAGTGAGGA 666
QY 659 TGTGCGGCTGCCAGCTGGGACTGCCCCACGCCCAAGAGAAATACAGGTGCCAGGAAAGTG 718
Db 667 TGTGCGGCTGCCAGCTGGGACTGCCCCACGCCCAAGAGAAATACAGGTGCCAGGAAAGTG 726
QY 719 CTGCCCCGAGTGGGTATGTGACCAGGGAGTGA---CACCGGCGATCCAGCGCTCCACGGC 775
Db 727 CTGCCCCGAGTGGGTGTGACCAGGCGAGTGATGACGCCGCAATCCAGCCCTCTCAGC 786
QY 776 GCAAGGACACCAACTTCTGCCCCTTGTCACTCCTGCCTCTGCTGATGCTCTTGTCCAAA 835
Db 787 CCAAGGACACCAACTTCTGCCCCTTGTCACTCCTGCATCTGCCGATGCGCCCTGTCCAAA 846
QY 836 TTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGT 895
Db 847 CTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGT 906
QY 896 GTCCAACCAAGAACCGATTCTGCCAATCGAGATCCAACGCCCTGTGTCTGCCAGACC 955
Db 907 ATCCAACCAAGAACCGATTCTGCCAATCGAGATCCAACGCCCTGTGTCTGCCAGACC 966
QY 956 CTGCCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCA-ACTGGGGA 1014
Db 967 CTGCCTGGCATCCAGGAGCCACGGCTCATGGAACAGTGCCTTCTAGAGCCATTCGGGGA 1026
QY 1015 TGGGATACAGGGCCTGCCATCCTCAGCAATGACCCCTAGGACCAGGCCCTGGACTGCTG 1074
Db 1027 TGTGGATACAGGGCCTGCCATTCACAGCAATGTCCCTAGGACCAGGCCCTGGACTGATG 1086
QY 1075 GTAGATGCTCTTCTCCATGCTCTTGGCTGCAGTTAACTGTCTCTTGGATTCACTGTGT 1134
Db 1087 GTAGATGCCCTCTCCATGCTCTTGGCTGCAGTTAACTGTCTCTTGGTGGATTCACTGTTC 1146
QY 1135 AGAGCCACTGAGCGATCCCTGCTCTGTGAGGTAGCGGAGCAGGTGACCACTCCAGT 1194
Db 1147 AGAGCCTCTGAGCGATCCCTGCTCTGTGAGGTGGGGAGCAGGTGACCACTCCATT 1206
QY 1195 TCTCTGG--TTCAGCCTGGAATTCTGGGTTCTCTGGCTCATTCCTCAAAACATCCCTGT 1252
Db 1207 TCTCTGGATTCTGACCCAGGCTTCTGGTTCTCTGGCTAGTTCTCTCAAAACTTCCCTGT 1266
QY 1253 ACAAAAAGGACAAACCAAAAGACCTTTAAACCTAGGCTATAGTGGGCAAACTGGCCACC 1312
Db 1267 ATGAAGAAGGACAAACCAAAAGGACCTTTAAAGCTAAGCTGTACTGGGCAAGCCTGGCCACC 1326

QY 1313 GTGCTGGGGATAGGTCAATGTTAGG-ACCAGACAGCAGATTGCCTGAAACTTCCAATTC 1371
|||||
Db 1327 ATGCTGGGGATAGTGACAGTAATAGGTACCAGGCAGCAGATTGCCTGAAACATCCAGGTC 1386

QY 1372 CCTTCTTGGACITCTGTATGCTTGTGTCCTCCCAAGATGATGAATGAACCTCCTAAGTGACCT 1431
|||||
Db 1387 CCTTCTTGGACITCTGTATGCTTGT-CCCAAGATTATGGGTGACCTTGTAAAGTGTGCCT 1445

QY 1432 TCCCTGACCTGAGAACACCCCTGCCTGCTCGGGAAGTATTCAGGGGGCAGAAATCTCTGTGA 1491
|||||
Db 1446 TTCCTGATCTGAGAACACCCCTGCCCGGCTGGGA-----AGAAATTTCTGGGA 1492

QY 1492 ACATGAAGAGAT-GAATCACACTGTCTTAAGCACTCGCAAAATTCCTGAAAGTCCAGGAACCTTGAGC 1550
|||||
Db 1493 ACATGAAGAGATGGAATCACACTATTCTTAAGAGCGTTTGCCAAAGTCCAGGAACCTTGACC 1552

QY 1551 TTTGTATTTTCAGGAATGCACATCTCTTAAAGCACTCGCAAAACAGGAAGGCTCCACACCT 1610
|||||
Db 1553 TTTGTATTTGTAATAATACACATCTCTTAAATGCTCACAAAGCAAG-AGGCTCCACACTT 1611

QY 1611 CTGGCAGGCCAGGSCCTTTCTCTTCAGCATGAGAAAGACAAGGGACAGCAGAGTACTCTC 1670
|||||
Db 1612 CTGGCAGGCCAGGSCCTTTCTCTTCAGCATGAGAGAGACAAGGAACAGTAGAGTACCCCTC 1671

QY 1671 CTCTGGAGGACTACTCTAGCCTAGAAATAAACACCCAAA 1708
|||||
Db 1672 CTCTGGAGGACTGSCCCCGTCTGGAATAAACACCCAAA 1709

RESULT 3
AAZ07517
ID AAZ07517 standard; cDNA; 753 BP.
XX AAZ07517;
XX
DT 26-NOV-1999 (first entry)
XX
DE Rat HICP polypeptide coding sequence.
XX
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX
OS Rattus sp.
XX
PN WO9947556-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Castellot JJ;
XX
DR WPI; 1999-562060/47.
DR P-PSDB; AAY27434.
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications -
XX
PS Claim 5; Fig 1; 108pp; English.
XX
CC The invention provides a rat heparin-induced CCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or

CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents the coding sequence of rat HICP.
XX
SQ Sequence 753 BP; 132 A; 235 C; 230 G; 156 T; 0 other;

Query Match 44.18; Score 753; DB 20; Length 753;
Best Local Similarity 100.0%; Pred. No. 3.8e-216;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCCTTCTCTCTCTCTCTCAATG 308
|||||
Db 1 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCCTTCTCTCTCTCTCTCAATG 60

QY 309 GTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACACCCAGTGCCCCA 368
|||||
Db 61 GTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACACCCAGTGCCCCA 120

QY 369 CAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACGGAGGCTG 428
|||||
Db 121 CAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACGGAGGCTG 180

QY 429 GGGAGTCTTGACACACCTGCATGTCTGCACCCAGCCAGGGCCTGGTTGTGAGCCT 488
|||||
Db 181 GGGAGTCTTGACACACCTGCATGTCTGCACCCAGCCAGGGCCTGGTTGTGAGCCT 240

QY 489 GGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 548
|||||
Db 241 GGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 300

QY 549 GAGTGATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTG 608
|||||
Db 301 GAGTGATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTG 360

QY 609 TGCCGCTGTATGACGGTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 668
Db 361 TGCCGCTGTATGACGGTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

QY 669 CCCAGCTGGACTGCCACGCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCCGAG 728
|||||
Db 421 CCCAGCTGGACTGCCACGCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCCGAG 480

QY 729 TGGGTATGTGACAGGGAGTGACACCCGGCGATCCAGCGCTCCACGGCGCAAGGACACCAA 788
|||||
Db 481 TGGGTATGTGACAGGGAGTGACACCCGGCGATCCAGCGCTCCACGGCGCAAGGACACCAA 540

QY 789 CTTTCTGCCCTTGTCACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 848
|||||
Db 541 CTTTCTGCCCTTGTCACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

QY 849 TGGGGCCCTGCTCAACCACTGTGGGTGGGCATAGCCACCGAGTGTCCAAACCCAGAAC 908
|||||
Db 601 TGGGGCCCTGCTCAACCACTGTGGGTGGGCATAGCCACCGAGTGTCCAAACCCAGAAC 660

QY 909 CGATTCTGCCAACTGGAGATCCAAACCGCGCTGTGTCTGCCCCAGACCCCTGGCAGCC 968
|||||
Db 661 CGATTCTGCCAACTGGAGATCCAAACCGCGCTGTGTCTGCCCCAGACCCCTGGCAGCC 720

QY 969 AGGAGCCACAGCTCATGGAACAGTGTCTTTCTAA 1001
|||||
Db 721 AGGAGCCACAGCTCATGGAACAGTGTCTTTCTAA 753

RESULT 4
AAZ07521
ID AAZ07521 standard; cDNA; 681 BP.
XX AAZ07521;
AC AAZ07521;
XX
DT 26-NOV-1999 (first entry)
XX
DE Rat HICP mature polypeptide coding sequence.
XX
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;

KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;
KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
KW gene therapy; transgenic animal; ss.

XX Homo sapiens.

XX WO200053753-A2.

XX PD 14-SEP-2000.

XX 05-JAN-2000; 2000WO-US00219.

XX 08-MAR-1999; 99WO-US05028.

XX 12-MAR-1999; 99US-0123957.

XX 14-MAY-1999; 99US-0134287.

XX 02-JUN-1999; 99WO-US12252.

XX 23-JUN-1999; 99US-0141037.

XX 20-JUL-1999; 99US-0144758.

XX 26-JUL-1999; 99US-0145698.

XX 01-SEP-1999; 99WO-US20111.

XX 08-SEP-1999; 99WO-US20594.

XX 15-SEP-1999; 99WO-US21090.

XX 15-SEP-1999; 99WO-US21547.

XX 05-OCT-1999; 99WO-US23089.

XX 30-NOV-1999; 99WO-US28313.

XX 30-NOV-1999; 99WO-US28409.

XX 02-DEC-1999; 99WO-US28564.

XX 02-DEC-1999; 99WO-US28565.

XX (GETH) GENENTECH-INC.

XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;

XX Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;

XX Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;

XX WPI; 2001-090793/10.

XX P-PSDB; AAB53084.

XX New isolated nucleic acid for producing a PRO polypeptide, analyzing
XX genetic disorders and treating cardiovascular, endothelial or
XX angiogenic disorders, such as atherosclerosis, wounds or cancer -

XX Claim 58; Fig 41; 293pp; English.

XX The invention relates to novel human angiogenesis-associated proteins
XX designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding
XX PRO proteins. The invention also relates to vectors and host cells
XX comprising a PRO nucleic acid, the recombinant production of a PRO
XX protein, PRO antibodies specific for a PRO protein, fusion proteins
XX comprising a PRO protein, agonists or antagonists of a PRO protein, and
XX compounds which inhibit the expression of a PRO gene. The invention
XX additionally encompasses methods of identifying modulators of PRO
XX expression or activity; diagnosing a cardiovascular, endothelial or
XX angiogenic disorder, or a susceptibility to such a disorder by detecting
XX mutations in a PRO gene, or the expression level of a PRO gene within a
XX particular tissue; treating a cardiovascular, endothelial or angiogenic
XX disorder via the administration of a PRO protein, PRO nucleic acid, or
XX PRO agonist or antagonist; a retroviral gene therapy vector comprising a
XX PRO nucleic acid; and methods of inhibiting or stimulating endothelial
XX cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the
XX administration of a PRO protein, or an agonist or antagonist thereof.
XX PRO nucleic acids, PRO proteins, antibodies against PRO proteins, PRO
XX agonists and PRO antagonists may be used as therapeutic agents to treat
XX cardiovascular, endothelial or angiogenic disorders, such as
XX atherosclerosis, osteoporosis, myocardial infarction, hypertension,
XX diabetic retinopathy, rheumatoid arthritis, Crohn's disease, psoriasis,
XX endometriosis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's
XX disease, or stroke. PRO nucleic acids are additionally useful in the
XX recombinant production of PRO proteins, as hybridisation probes to
XX screen libraries to isolate cDNAs with sequence identity to PRO proteins,
XX to map genes encoding PRO proteins, to analyse genetic disorders, and in
XX gene therapy. PRO nucleic acids can also be used to produce transgenic

CC animals useful for the development and screening of potential
CC therapeutic agents. The present sequence represents a cDNA encoding a PRO
CC protein of the invention.

XX

SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match

Best Local Similarity 32.9%; Score 561.4; DB 22; Length 1266;

Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCCTTCTCCTTCTC 302

Db 4 GGGGACATGAGAGGCACACCGAAGACCCACCTCTTCCCTTCTCCTTCTCCTCCTC 63

QY 303 TCAATGGTGTGTGCCAGCTGTGCCGACACACCTGTACCTGTCTTGGACACACCCAG 362

Db 64 TCAAAGGTGGTACCCAGCTGTGCCGACACACCTGTACCTGTCTTGGACACCTCCCGA 123

QY 363 TGCCACACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 422

Db 124 TGCCCGCTGGGAGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 183

QY 423 AGGCTGGGGAGTCTCGGACACCTGCATGTCTGCGACCCAGCCAGGCGCTGGTTGT 482

Db 184 CGGCTGGGGAGTCTCGGACCACTCCAGCTCTGCGACCCAGCCAGGCGCTGGTCTGC 243

QY 483 CAGCCTGGGGAGGCGCTGGCGGCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT 542

Db 244 CAGCCCGGGGAGGACCCGGTGGCGGGGGGCGCTGTGCCTCTTGGCAGAGGACGACAGC 303

QY 543 AGCTGTGAGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602

Db 304 AGCTGTGAGTGAACGGCCGCTGTATCGGAAGGGAGACCTTCCAGCCCCACTGCAGC 363

QY 603 GTCCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGCAGTGAGGATGTG 662

Db 364 ATCCGCTGCCGCTGGGAGGACGGCGGCTTCACTGCGCTGTGCAGGAGGATGTG 423

QY 663 CGGCTGCCAGCTGGGACTGCCACGCCCAAGAGATAACAGTGGCCAGGAAAGTGTGC 722

Db 424 CGGCTGCCAGCTGGGACTGCCCCACCCAGGAGGTCGAGGTCTTGGCAAGTGTGC 483

QY 723 CCGAGTGGGTATGTGACCAGGAGTGACACCGCGCATCCAGCGCTCCACGGCGCAAGGA 782

Db 484 CCTGAGTGGGTGTGGGCCAAGGAGGGGACTGGGACCCAGCCCTTCCAGCCCAAGGA 543

QY 783 CACCAACTTCTGCCCTTGTCACTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 842

Db 544 CCCCAGTTTCTGGGCTTGTCTTCCCTGCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 603

QY 843 ACAGCCTGGGCGCCCTGTCTCAACACCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 902

Db 604 ACGGCTGGGACCTGTCTCGACCACTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 663

QY 903 CAGAACCGATTCTGCCAACTGGAGATCCAAACGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 962

Db 664 CAGAACCGCTTCTGCCGACTGGAGACCCAGCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 723

QY 963 GCAGCCAGGAGCCACAGCTCATGGACAGTGTCTTCTA-AGGCCAACTGGGATGCGGAT 1021

Db 724 CCTCCAGGGTGGCAGTCCAAACACAGTGCCTTCTAGAGCCGGGCTGGGAATGGGAC 783

QY 1022 ACAGGCTGCCATCTCTCAGCAATGACCCCTAGGACCCCTAGGACCCCTAGGACCCCTAGGATG 1081

Db 784 ACGGTGTCCACCATCCCAAGTGGTGGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 843

QY 1082 CT-CTTCTCCATGCTTGGCTGCAGTAACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1139

Db 844 GTCCGTGCCAGGCGCTTGGCTGCAGGCAACACTTTAGCTTGGTGGTCCACCATGCAAGAAC 902

RESULT 13
AAV29260

ID AAV29260 standard; DNA; 1285 BP.
XX
AC AAV29260;
XX
DT 14-SEP-1998 (first entry)
XX
DE Human connective tissue growth factor-3 gene.
XX
KW Connective tissue growth factor-3; CTGF-3; human; cancer;
KW arthritis; fibrosis; osteoporosis; diagnosis; therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 9..761 /*tag= a
FT sig_peptide 9..65 /*tag= b
FT mat_peptide 66..758 /*tag= c
FT
XX
PN WO9821236-A1.
XX
PD 22-MAY-1998.
XX
PF 08-NOV-1996; 96WO-US17856.
XX
PR 08-NOV-1996; 96WO-US17856.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Chopra A, Ebner R, Ruben SM;
XX
XX WPI; 1998-297864/26.
DR P-PSDB; AAW37946.
XX
PT Novel human connective tissue growth factor 3 gene - useful for the
PT diagnosis and treatment of e.g. cancer, arthritis, fibrosis,
PT osteoporosis
XX
PS Claim 2; Fig 1A-B; 87pp; English.
XX
CC This nucleotide sequence codes for human connective tissue growth
CC factor-3 (CTGF-3) protein (see AAW37946), a novel member of the
CC growth factor superfamily. It was discovered in a cDNA library
CC derived from human osteoblasts. The gene has also been identified
CC in cDNA libraries from ovary, testis, heart, lung, skeletal muscle,
CC adrenal medulla, adrenal cortex, thymus, prostate, small intestine
CC and colon. A cDNA clone is deposited as ATCC 97756. Also provided
CC are vectors, host cells and recombinant methods for producing
CC CTGF-3 polypeptides. CTGF-3 nucleic acid sequences or their
CC fragments, e.g. primers or probes, can be used to diagnose diseases
CC where CTGF-3 expression is enhanced, e.g. cancer, arthritis,
CC fibrosis or atherosclerosis, or diseases where expression is
CC decreased such as in osteoporosis. Disorders characterised by
CC decreased or increased levels of CTGF-3 can be treated by
CC administering CTGF-3 polypeptides and anti-CTGF-3 antibodies,
CC respectively.
XX
SQ Sequence 1285 BP; 237 A; 418 C; 389 G; 241 T; 0 other;

Query Match 32.9%; Score 561.4; DB 19; Length 1285;
Best Local Similarity 78.0%; Pred. No. 2.7e-158;
Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

QY 243 GGTGACATGAGGGGAGCCACCTGATCCATCTTCTGGCCACTCCTTCTCTGCTTC 302
|| ||||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 GGGACATGAGAGGCACACCGAAGACCCACCTCTGGCCTTCTCCCTCTCTGCTCCTC 62

QY 303 TCAATGTTGTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCCACCCAG 362
||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db TCAAAGGTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCCCTGGCCACCTCCCCGA 122

QY 363 TGCCACAGGGGGTACCCCTGGTGGATGGCTGTGGCTGTAAAGTGTGTGCACGG 422
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123 TGCCCGCTGGGAGTACCCCTGGTGGATGGCTGTGGCTGTCCGGGTATGTGCACGG 182

QY 423 AGGCTGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGGCGCTGGTTGT 482
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 183 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCGCAGGCGCTGGTCTGC 242

QY 483 CAGCTGGGGCAGGCCCTGGCGCATGGGGCTGTGTGTCTCTGGATGAGGATGACGGT 542
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 CAGCCCGGGCAGGACCCGCTGCCGGGGGCCCTGTGCTCTTGGCAGAGGACGACAGC 302

QY 543 AGCTGTAGGTGAATGGCCGACGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 AGCTGTAGGTGAACGGCCGCCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 362

QY 603 GTCCTGTCCCGCTGTGATGACGGTGGCTTCACCTGCCTGCCCTGTGCAGTGAGGATGTG 662
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 363 ATCCGCTGCCGCTGCGAGGACGGCGCTTCACCTGCGTGGCTGTGCAGCGGAGGATGTG 422

QY 663 CGGCTGCCCAGCTGGGACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGC 722
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 423 CGGCTGCCCAGCTGGGACTGCCCCACCCAGGAGGGTGCAGGTCTTGGCAAGTGTGC 482

QY 723 CCCGAGTGGGTATGTGACCAAGGAGTGACACCGCGCATCCACGGCTCCACGGCGCAAGGA 782
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 483 CCTGAGTGGGTGTGCGGCCCAAGGAGGGGACTGGGACCCAGCCCCCTTCCAGCCCCAAGGA 542

QY 783 CACCAACTTCTGCCCTTGTCACTCCTCCTCTGCTGATGCTCTTGTCCAAATGGAGC 842
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 543 CCCAGTTTCTGGCCTTGTCTCTTCCCTGCCCTGGTGTCCCTGCCAGAATGGAGC 602

QY 843 ACAGCCTGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAC 902
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 603 ACGGCTGGGACCCCTGCTCGACCACTGTGGGCTGGGCATGGCCACCCGGGTGTCCAAC 662

QY 903 CAGAACCGATTCTGCCAACTGGAGATCCAACGCCCTGTGTCTGTGCTGCCAGACCCCTGCCTG 962
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 663 CAGAACCGCTTCTGCCGACTGGAGACCCAGCCGCCCTGTGCTGTCCAGGCCCTGCCCA 722

QY 963 GCAGCCAGGAGCCACAGCTCATGGAACAGTCTTTCTA-AGGCCAACTGGGATGCGGAT 1021
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 723 CCCTCCAGGGGTGCGAGTCCACAAAACAGTGCCTTCTAGAGCGGGCTGGGAATGGGGAC 782

QY 1022 ACAGGCGCTGCCATCCTCAGCAATGACCCCTAGGACCAAGCCCTGGACTGTGGTAGATG 1081
|| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 783 ACGGTGTCCACCATCCCCAGCTGGTGGCCCTGTGCTTGGGCCCTGGGCTGATGGAAGATG 842

QY 1082 CT-CTTCTCCATGCTCTTGGCTGCAGTTAACTGTCTGCTTGGATTCACTGTGTAGAGC 1139
|| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 843 GTCGCTGCCCAGGCCCTTGGCTGCAGGCAACACTTTAGCTTGGTCCACCATGCAGAAC 901

RESULT 14
AAx76486
ID AAX76486 standard; DNA; 1293 BP.
XX
AC AAX76486;
XX
DT 06-AUG-1999 (first entry)
XX
DE Human WISP-2 protein nucleotide sequence SEQ ID NO:13.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
OS Homo sapiens.
XX

PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI; 1999-337420/28.
DR P-PSDB; AAY17649.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Example 4; Page 174-175; 284pp; English.
XX
CC The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 1293 BP; 232 A; 425 C; 393 G; 243 T; 0 other;

Query Match	32.9%;	Score 561.4;	DB 20;	Length 1293;
Best Local Similarity	78.0%;	Pred. No. 2.7e-158;		
Matches 701;	Conservative 0;	Mismatches 196;	Indels 2;	Gaps 2
QY	243	GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCCTCTCCCTTCTC	302	
Db	16	GGGGACATGAGAGGCACACCGAAGACCCACCTCCTGGCCCTCTCTCCCTCCTCTGCCCTCCTC	75	
QY	303	TCAATGGTGTGTGCCAGCTGTGTCGGACACCCTGTACCTGTCTCTTGGACACCACCCACG	362	
Db	76	TCAAAGGTGCGTACCAGCTGTGCCCCGACACCATGTACCTGCCCTGGCCACCTCCCCGA	135	
QY	363	TGCCCCACAGGGGGTACCCCTGGTCTGGATGGCTGTGGCTGTCTGTAAAGTGTGTGCACGG	422	
Db	136	TGCCCCGTGGGAGTACCCCTGGTCTGGATGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGC	195	
QY	423	AGGCTGGGGGAGTCTTCGACACCACCTGCATGTCTGCGACCCCAAGCCAGGCGCTGGTTGT	482	
Db	196	CGGCTGGGGGAGCCCTGCGACCAACTCCACGTCTGCGAGCCAGCCAGGCGCTGGTCTGC	255	
QY	483	CAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTTGGATGAGGATGACGGT	542	
Db	256	CAGCCCGGGGCAGGACCCGGTGGCCGGGGCCCTGTGCCCTCTTGGCAGAGGACGACAGC	315	
QY	543	AGCTGTGAGGTGAATGGCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG	602	
Db	316	AGCTGTGAGGTGAACGGCCCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCCACTGCAGC	375	

QY	603	GTCCTGTGCCGCTGTGATGACGGTGGCTTTCACTGCCTGCCCGCTGTGCAGTGAGGATGTG	662
Dd	376	ATCCGCTGCCGCTGCGAGGACGGCGGCTTACCTGCGTGCCGCTGTGCAGCAGGAGTGTG	435
QY	663	CGGCTGCCCCAGCTGGGACTGCCCACGCCCCCAGAGAATAACAGGTGCCAGGAAGTGC	722
Dd	436	CGGCTGCCCCAGCTGGGACTGCCCCCACCAGGAGGGTCGAGGTCTTGGSCAAGTGC	495
QY	723	CCCGAGTGGGTATGTGACCAAGGGAGTGACACCGGGCGATCCAGCGCTCCACGGCGCAAGGA	782
Dd	496	CCTGAGTGGGTGTGCGGCCAACGAGAGGGGAGTGGGACCCAGCCCCCTTCACGCCCAAGGA	555
QY	783	CACCAAACTTCTGGCCCTTGTCACTCCTGCCTCTGCTGATGCTCCTTGTCTCCAATTGGAGC	842
Dd	556	CCCCAGTTTTCTGGCCCTTGTCTCTTCCCTGCCCTTGGTGTCCCTGCCCAAGATGGAGC	615
QY	843	ACAGCCTGGGCCCCCTTGCTCAAACCACTGTGGGCTGGGCATAGCCACCCAGTGTCCAAC	902
Dd	616	ACGGCCTGGGACCCCTGCTCGACCACCTGTGGGCTGGGCATGGCCACCCCGGTGTCCAAC	675
QY	903	CAGAACCGATTGTGCCAACTGGAGATCCCAACGCCCGCCTGTGTCTGCCACAGACCCCTGC	962
Dd	676	CAGAACCGTTGTGCCGACTGGAGACCCACGCGCCGCTGTGCCTGTCCAGGCCCTGCCCA	735
QY	963	GCAGCCAGGAGCACAGCTCATGGAAACAGTCTTTCTA-AGGCCAACTGGSGATGCGGAT	1021
Dd	736	CCCTCCAGGGTCCGAGTCCACAACACAGTGCCTTCTAGAGCCGGCTGGGAATGGGAC	795
QY	1022	ACAGGCGCTGCCATCCTCAGCAAAATGACCTAGGACAGGCCCTGGACTGCTGTAGATG	1081
Dd	796	ACGGTGTCCACCATCCCACTGGTGGCCCTGTGCCTGGGCCCTGGGCTGATGGAAGATG	855
QY	1082	CT-CTTCTCCATGCTTTGGCTGCAGTTAACTGTCTCTTGGATTCACTGTGTAGAGC	1139
Dd	856	GTCCGTCGCCAGGCCCTTGCTGCAGGCAACACTTAGCTTGGTCCACCATGCAGAAC	914
 RESULT 15 AAH28214			
XX	ID	AAH28214 standard; cDNA; 1309 BP.	
AC	AAH28214;		
XX	OS	Homo sapiens.	
XX	DE	Connective tissue derived growth factor related protein cDNA.	
XX	KW	Growth factor; protein inhibitor; protease; damaged tissue;	
XX	KW	platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;	
XX	KW	connective tissue derived growth factor; CTGF; chrysalin; VEGF;	
XX	KW	keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;	
XX	KW	transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;	
XX	KW	granulocyte macrophage colony stimulating factor; GM-CSF; uPA;	
XX	KW	vascular endothelial growth factor; urokinase plasminogen activator;	
XX	KW	dermal ulcer; wound; ss.	
OS	Key	Location/Qualifiers	
XX	FH	7..759	
FT	CDS	/tag=a	
FT		/product= "connective tissue derived growth factor	
FT		related protein"	
XX	PN	WO200149309-A2.	
XX	PD	12-JUL-2001.	
XX	PD	21-DEC-2000; 2000WO-IB01935.	
PF		29-DEC-1999; 99GB-0030768.	
XX	PR	(PFI2) PFIZER LTD.	
XX	PA		

PA (PFIZ) PFIZER INC.
XX
PI Davies MJ, Huggins JP, McIntosh FS, Occleston NL;
XX
DR WPI; 2001-418351/44.
DR P-PSDB; AAB84599.
XX
PT Composition for the treatment of damaged tissue i.e. chronic wounds and
PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
PT factor -
XX
XX
PS Disclosure; Page 546; 572pp; English.
XX
CC The specification describes a pharmaceutical composition, comprising
CC a growth factor, an inhibitor agent, i.e. a protease. The inhibitor
CC agent inhibits the action of at least one specific adverse protein,
CC i.e. a protease, that is upregulated in a damaged tissue such as a
CC wound environment. Growth factors which are included in the composition
CC of the invention are platelet-derived growth factor (PDGF), fibroblast
CC growth factor (FGF), connective tissue derived growth factor (CTGF),
CC keratinocyte-derived growth factor (KGF), transforming growth
CC factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor
CC (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth
CC factor (VEGF), and chrysalin. Inhibitors which are included in the
CC composition of the invention include inhibitors of urokinase-type
CC plasminogen activator (uPA) and matrix metalloproteinase (MMP). The
CC composition is useful for the treatment of chronic damaged tissue, i.e.
CC wounds and dermal ulcers. The present sequence encodes a human
CC CTGF-like protein, and is used to produce the composition of the
CC invention.
XX
SQ Sequence 1309 BP; 261 A; 418 C; 387 G; 242 T; 1 other;

Query Match 32.9%; Score 561.4; DB 22; Length 1309;
Best Local Similarity 78.0%; Pred. No. 2.7e-158;
Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

QY 243 GGTGACATGAGGGGAGCCGACCTGATCCACTTCTTGGCACCTTCTTCCCTGCTCTC 302
Db 1 GGGGACATGAGAGGCACACCGAAGACCCACCTCTGCTGCTCTTCCCTCTGCTCTC 60

QY 303 TCAATGGTGTGTGCCAGCTGTGCGGACACCCCTGTACTGTCTTGGACACCCACCCAG 362
Db 61 TCAAAGGTGCGTACCCAGCTGTGCCCCACACCATGTACTGTCCCTGGCCACCTCCCCG 120

QY 363 TGCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTAAAGTGTGACAGG 422
Db 121 TGCCGCTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTGCGGGTATGACAGG 180

QY 423 AGCTGGGGGAGTCCCTGCACCACTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 482
Db 181 CGCTGGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGCGCTGCTGC 240

QY 483 CAGCCTGGGGCAGGCCCTGGGGCCCATGGGGCTGTGCTGTCTTGGATGAGGATGACGGT 542
Db 241 CAGCCGGGGCAGGACCCGCTGGCGGGGGCCCTGTGCTGTGCTGTGCTGTGCTGTG 300

QY 543 AGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGACACCTTTAAACCCCAATTGCAGG 602
Db 301 AGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 360

QY 603 GTCCTGTGCCGCTGTGATGACGGTGGCTTCACTGCTGCCGCTGTGCTGTGCTGTGCTGTG 662
Db 361 ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACTGCTGCCGCTGTGCTGTGCTGTGCTGTG 420

QY 663 CGGCTGCCCCAGCTGGGACTGCCACGCCCCCAAGAGATACAGGTGCCAGGAAAGTGTCTGC 722
Db 421 CGGCTGCCCCAGCTGGGACTGCCACGCCCCCAAGAGATACAGGTGCCAGGAAAGTGTCTGC 480

QY 723 CCCGAGTGGGTATGTACAGGGGAGTGACACCGGGCATCCAGCGCTCCACGGCGCAAGGA 782
Db 481 CCTGAGTGGGTGTGCGGCCCAAGGAGGGGGACTGGGGACCCAGCCCCCTTCCAGCCCCAAGGA 540

Search completed: July 23, 2003, 11:15:15
Job time : 455.287 secs

QY 783 CACCAACTTTCTGCCCCCTGTGTCACTCCCTCTCTGTGTGATGCTCCCTTCCAAATGGAGC 842
Db 541 CCCAGATTTCTGGCCTTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCAGATGGAGC 600

QY 843 ACAGCCTGGGGCCCCCTGTCAACCAACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAC 902
Db 601 ACGGCCTGGGACCCCTGTCTGACCACTGTGGGCTGGGCATGGCCACCCGGGTGTCCAAC 660

QY 903 CAGAACCGATTCTGCCAACTGGAGATCCAACGCGCCTGTGTCTGTGCCCCAGACCCCTGCTG 962
Db 661 CAGAACCGCTTCTGCCGACTGGAGACCCAGCGCGCCTGTGCTGTGCCCCCTGCCCA 720

QY 963 GCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTTCTA-AGGCCAACTGGGGATGGGAT 1021
Db 721 CCTCCAGGGGTGCGAGTCCACAAAACAGTGCCTTTAGAGCCGGGCTGGGAATGGGGAC 780

QY 1022 ACAGGGCCTGCCATCTCAGCAAATGACCCCTAGGACCCAGGCCCCCTGGACTGCTGGTAGATG 1081
Db 781 ACGGTGTCCACCATCCCAAGCTGGTGGCCCTGTGCTGGGCCCTGGGCTGATGGAGATG 840

QY 1082 CT-CTTCTCCATGCTTGGCTGCAGTTAACTGTCTGCTTGGATTCACTGTGTAGAGC 1139
Db 841 GTCCGTGCCAGGCCCTTGGCTGCAGGCAACACTTTAGCTTGGGTCCACCATGCAGAAC 899

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:55:06 ; Search time 92.9244 Seconds
(without alignments)
5636.882 Million cell updates/sec

Title: US-10-010-408-1
Perfect score: 1708
Sequence: 1 GACGGCTTCTGATCTCCAGAG.....GCCTAGATAAACACCCAAA 1708

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1278	74.8	1734	4	US-09-182-145-17
c	2	1278	1734	4	US-09-182-145-18
	3	561.4	32.9	4	US-09-182-145-13
c	4	561.4	32.9	4	US-09-182-145-14
	5	528.8	31.0	4	US-09-182-145-39
6	501.2	29.3	738	4	US-09-182-145-38
7	163.8	9.6	2075	1	US-08-167-628-1
8	163.8	9.6	2075	1	US-08-386-680-1
9	163.8	9.6	2075	1	US-08-459-717-1
10	163.8	9.6	2075	1	US-08-712-302-1
11	163.8	9.6	2075	2	US-08-880-031-1
12	163.8	9.6	2075	3	US-09-097-179-1
13	163.8	9.6	2075	4	US-09-080-715-1
14	163.8	9.6	2075	4	US-09-142-569-7
15	163.8	9.6	2075	5	PCT-US96-08140-1
16	163.8	9.6	2998	3	US-09-054-368-1
17	163.8	9.6	2998	3	US-09-054-274-1
18	163.8	9.6	2998	4	US-09-056-704-1
19	161	9.4	2267	4	US-09-142-569-5
20	156.4	9.2	2350	4	US-09-187-478-1
21	154.8	9.1	2350	4	US-09-292-036-1
22	123.4	7.2	1418	4	US-09-142-569-3
23	119.8	7.0	1766	4	US-09-182-145-9
c	24	119.8	7.0	4	US-09-182-145-10
	25	117	6.9	4	US-09-142-569-1
26	111.6	6.5	2830	4	US-09-182-145-1
c	27	111.6	6.5	4	US-09-182-145-2

28	102.8	6.0	1128	2	US-08-459-101A-1	Sequence 1, Appli
29	98.4	5.8	1062	4	US-09-253-316-3	Sequence 3, Appli
30	89.2	5.2	1403	4	US-09-182-145-23	Sequence 23, Appl
31	84	4.9	4214	4	US-09-122-135-1	Sequence 1, Appli
32	75.4	4.4	1142	4	US-09-253-316-1	Sequence 1, Appli
33	75.4	4.4	1212	4	US-09-182-145-34	Sequence 34, Appl
34	75.4	4.4	1212	4	US-09-182-145-35	Sequence 35, Appl
35	75.4	4.4	1335	4	US-09-182-145-30	Sequence 30, Appl
36	75.4	4.4	1335	4	US-09-182-145-31	Sequence 31, Appl
37	65.8	3.9	1101	4	US-09-182-145-29	Sequence 29, Appl
38	63.4	3.7	693	4	US-09-182-145-24	Sequence 24, Appl
39	63.4	3.7	1202	4	US-09-182-145-26	Sequence 26, Appl
40	48.6	2.8	2541	2	US-08-656-393-1	Sequence 1, Appli
41	41.4	2.4	51	4	US-09-182-145-117	Sequence 117, App
42	39.8	2.3	546	4	US-09-182-145-28	Sequence 28, Appl
43	39.8	2.3	683	4	US-09-182-145-25	Sequence 25, Appl
44	39.8	2.3	1183	4	US-09-182-145-27	Sequence 27, Appl
45	36.6	2.1	4403765	4	US-09-103-840A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-182-145-17
; Sequence 17, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-182-145-17

Query Match 74.8%; Score 1278; DB 4; Length 1734;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 1520; Conservative 0; Mismatches 165; Indels 33; Gaps 11;

QY	3	CGCTTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCCCTTGGCAAGGCTGGCAGCCGC	62
Db	13	CGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCCCTTGGCGAGGCTGCAGCTGC	72
QY	63	TG-GGCAGTGGCTTGAATGGAGGTCTTTATTACTGGAACTGAGGAGCTAAGAGGCTCC	121
Db	73	TGTGGCAGTAGCTTGGGATGGAGGTCTTTCTTGTGGAACTGAGGAGCTGAGAGGCTCC	132
QY	122	TGTCAG---CTTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTACACACTGTCA	178
Db	133	TGTCAGGCTCCTGTCTCTAAACTTCTTGGCACTTGGCGTGGCTTGGGCTTACACACTGTCA	192
QY	179	GACACCTTCGTGGTGGCTCCACGGCCTCACCTTCAGGTTGAAGCTGGCTCCACAAGGG	238

QY	122	TGTCAG-- -CTGTCTCCTAAAGTCTTAGCACATTGTGGTGGCTTTGGGCTTCACACACTGTCA	178
Db	1602	TGTCAGGCTCCTGTCTTAACCTCTTGGCACATTGCGGTGGCTTGGGCTTCACACACTGTCA	1543
QY	179	GACACCTTCGTGGTGGCCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG	238
Db	1542	GACACCTTCGTGGTGGCCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG	1489
QY	239	ACACGGTGACATGAGGGGCAGCCCACCTGATCCATCTTCTGGCCACTTCTTCCTCTCTGTCCT	298
Db	1488	ACACGGTGACATGAGGGGCACCCACTGATCCATCTTCTGGCCACTTCTTCCTCTCTGTCAT	1429
QY	299	TCTCTCAATGGTGTGTGCCAGAGTGTGCCGGACACCCTGTACCTGTCTGGCCACTTCTTCCTCTCTGTC	358
Db	1428	TCTCTCAATGGTGTATTCCAGAGTGTGCCAGCACCCCTGTGCCTGTCTTCCTTGGACACCACC	1369
QY	359	CCAGTGCCCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTTGAAGTGTGTGC	418
Db	1368	CCAGTGCCCCACGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTCGAGTGTGTGC	1309
QY	419	ACGGAGGCTGGGGGAGTCTTGCAGACCACCTGSCATGTCTGCGACCCCGCCAGCCAGGGCCCTGGT	478
Db	1308	ACGGAGGCTGGGGGAGTCTTGCAGACCACCTGSCATGTCTGCGACCCCGCCAGCCAGGGCCCTGGT	1249
QY	479	TTGTACGCCCTGGSGCAGGCCCTGGCGGCCCATGGGGCTGTGTGTCTCTTGGATGAGGATGA	538
Db	1248	TTGTACGCCCTGGSGCAGGCCCCAGTGGCCGTGGTGTGTGTGCCTCTTTCGAAGAGGATGA	1189
QY	539	CGGTAGCTGTAGGTGAATGGCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAATTG	598
Db	1188	CGGGAGCTGTAGGTGAATGGCCGCAGGTACCTGGATGGGAGACCTTTAAACCCAATTG	1129
QY	599	CAGGTCCTGTGCCGCTGTGATGACGGTGGCTTCACTTCACTTGCCTGCCGTGTGTCAGTGAGGA	658
Db	1128	CAGGGTTTTGTGCCGCTGTGATGACGGTGGTTTCACTTGCCTGCCGTGCCGTGTGTCAGTGAGGA	1069
QY	659	TGTGCGGCTGCCAGCTGGGACTGCCACGCCCCAAAGAGAATAACAGGTGCCAGGAAAAGTG	718
Db	1068	TGTGCGGCTGCCAGCTGGGACTGCCACGCCCCAGGAGAATAACAGGTGCCAGGAAAGGTG	1009
QY	719	CTGCCCCGAGTGGGTATGTGACCAGGGAGTG- - -CACCGCGATCCAGCGCTCCACGGC	775
Db	1008	CTGCCCCGAGTGGGTGTGTGACCAGGCAGTGATGCAGCCCGCAATCCAGCCCTCCTCAGC	949
QY	776	GCAAGGACACCAAACCTTTCTGCCCTTGTCACTCTCTGCCTCTGTCTGATGCTCCTTGTCCAAA	835
Db	948	CCAAGGACACCAAACCTTTCTGCCCTTGTCACTCTCTGCATCTGCCGATGGCCCTGTCCAAA	889
QY	836	TTGGAGCACAGCCTGGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGT	895
Db	888	CTGGAGCACAGCCTGGGGCCCCCTGCTCAACCACCTGTGGGTGGGCTAGCCACCCGAGT	829
QY	896	GTCCAACCAAGAACCGATTCTTGCCAACTGGAGATCCAAACGCCGCTGTGTCTGCTGCCAGACC	955
Db	828	ATCCAACCAAGAACCGATTCTTGCCAACTGGAGATCCAGCGTGCCTGTGTCTGCTGCCAGACC	769
QY	956	CTGCCCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCA- ACTGGGGA	1014
Db	768	CTGCCCTGGCATCCAGGAGCCACGGCTCATGGAACAGTGCCTTCTAGAGCCATTGCGGGGA	709
QY	1015	TGCGGATACAGGGCCTGCCATCCTCAGCAATGACCCCTAGGACCAGGCCCTGGACTGCTG	1074
Db	708	TGTGGATACAGGGCCTGCCATTCTCAGCAATGTCCCTAGGACCAGGCCCTGGACTGATG	649
QY	1075	GTAGATGCTCTTCTCCATGCTCTTGGCTGCAGTTAACTGTCTCTTGGATTCACTGTGT	1134
Db	648	GTAGATGCCCTCTCCATGCTCTTGGCTGCAGTTAACTGTCTCTGGGTGGATTCACTGTCC	589
QY	1135	AGAGCCACTGACGATCCCTGCTCTGTCTGAGGTAGCGGAGCAGGTGACCAGCTCCAGT	1194
Db	588	AGAGCCCTGAGCGGATCCCTGCTCTGTCTGAGGTGGGGGAAGCAGGTGACAGCTCCATT	529

QY	1195	TCCTCTGG--TTACAGCCTGGAATTCCTGGGTCTCTCTGGCTCATTCCTCAAAACATCCCTGT	1252
Db	528	TCCTCTGGATTCTGACCCAGGCTTCCTGGGTTCTCTGGCTAGTTCCTCAAAACTTCCCTGT	469
QY	1253	ACAAAAAGGACAACCAAAAAAGACCTTTAAACCTAGGCTATACTGGGCAAACTGGGCCACC	1312
Db	468	ATGAAAAGGACAACCAAAAGGACCTTTAAAGCTAAGCTGTACTGGGCAAGCCTGGCCACC	409
QY	1313	GTGCTGGGGATAAGGTCAATGTTAGG-ACCAGACAGCAGATTGCCCTGAAACTTCCAATTC	1371
Db	408	ATGCTGGGGATAGTGACAGTAATAGGTACCAAGCAGCAGATTGCCCTGAAACATCCAGGTC	349
QY	1372	CCTTCTTGGACTTCTGTATGCTTGTCTCCCAAGAGATGATGAATGAACTCGTAAGTGTACCT	1431
Db	348	CCTTCTTGGACTTCTATGTCTTGT-CCCAAGATATATGGTGACCTTGTAGTGTGCCT	290
QY	1432	TCCCTGACCTGAGAACACCCCTGCCTGCTCGGGAAGTATTTCAGGGGCAGAAATCTCTGTGA	1491
Db	289	TTCTCTGATCTGAGAACACCCCTGCCCGGCTGGGA-----AGAAATTTCTTGGGA	243
QY	1492	ACATGAAGAGAT-GAATCACACTGTCTTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAGC	1550
Db	242	ACATGAAGAGATGGAATCACACTATTCTTTAAGAGCGTTTGCCAAAGTCCAGGAACCTTGACC	183
QY	1551	TTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCT	1610
Db	182	TTTGTATTTGTAAAAATACACATCTCTTTAAATGCTCACAAAGCAAG-AGGCTCCACACTT	124
QY	1611	CTGGCAGGCCAGGGCCCTTCTCTTCAGCATGAGAAAGACAAAGGACAGCAGTACTCTC	1670
Db	123	CTGGCAGGCCAGGGCCCTTCTCTTCAGCATGAGAGAGACAAAGGAACAGTAGAGTACCCCTC	64
QY	1671	CTCTGGAGGACTAGTCTAGCCTAGAAATAAACACCCAAA	1708
Db	63	CTCTGGAGGACTGGCCCGGCTCTGGAATAAACACCCAAA	26

RESULT 3

```

US-09-182-145-13
; Sequence 13, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 13
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-13

```

Query-Match 32.9%; Score 561.4; DB 4; Length 1293;
Best Local Similarity 78.0%; Pred. No. 2 2e-152;
Matches 701; Conservative 0; Mismatches 196; Indels 2;
Gaps 2;

WILSON, WILLIAM I. : TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

QY	243	GGTGACATGAGGGGACGCCACTGATCCATCTCTTGGCCACTTCCTTCCCTCTGCCTTCTC	302
Db	16	GGGGACATGAGAGGCACACCGAAGACCCACCTCTTGGCCTTCTCCCTCCTCTGCCTCTC	75
QY	303	TCAATGGTGTGTGCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCACG	362
Db	76	TCAAAGGTGCTACCCAGCTGTGCCCGACACCCATGTACCTGCCCTGGCCACCTCCCCGA	135
QY	363	TGCCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAAGTGTGTGCACGG	422
Db	136	TGCCCCGTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGCCGGGTATGTGCACGG	195
QY	423	AGGCTGGGGAGTCTTGGACCACTGCATGTCTGGACCCACCCAGCCAGGGCCTGGTTGT	482
Db	196	CGGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGC	255
QY	483	CAGCCTGGGGACGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGT	542
Db	256	CAGCCCGGGCAGAACCCGGTGGCCGGGGCCCTGTGTCTCTTGGCAGAGGACGACAGC	315
QY	543	AGCTGTAGGTGAATGGCCGACAGTACCTGGATGAGAGAGACCTTTAAACCCCAATTGCAGG	602
Db	316	AGCTGTAGGTGAACGGCCGCCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC	375
QY	603	GTCCGTGTCCCGCTGTGATGACGGTGGCTTCACCTGCCTGCCGCTGTGCAGTGAGGATGTG	662
Db	376	ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACCTGCGTGCCGCTGTGCAGCGGAGGATGTG	435
QY	663	CGGCTGCCCCAGCTGGGACTGCCCAAGCCCCAACAGAAATACAGGTGCCAGGAAAGTGCTGC	722
Db	436	CGGCTGCCCCAGCTGGGACTGCCCCCAACCCACGAGGGTCTGAGGTCTTGGGCAAGTGCTGC	495
QY	723	CCCGAGTGGGTATGTACCAGGGAGTGACACCGGGCATCCAGCGCTCCACGGCGCAAGGA	782
Db	496	CCTGAGTGGGTGTCCGGCCAAGGAGGGGACTGGGGACCCAGCCCCCTTCCAGCCCAAGGA	555
QY	783	CACCAACTTTCTGCCCCTGTCTACTCTCTGCCCTGTGCTGATGCTCTTGTCCAAATGGAGC	842
Db	556	CCCCAGTTTCTGSCCTGTCTCTTCCCTGCCCCCTGGTGTCCCTTGCACAGATGGAGC	615
QY	843	ACAGCCTGGGGCCCTGTCTCAACCACCTGTGGGTGGGCATAGCCACCCGAGTGTCCAAC	902
Db	616	ACGGCCTGGGGACCTGTCTGACCACCTGTGGGTGGGCATGGCCACCCGGGTGTCCAAC	675
QY	903	CAGAACCGATTCTGCCAACTGGAGATCCAACGCCCGCCTGTGTCTGCCACAGACCCCTGCCTG	962
Db	676	CAGAACCGCTTCTGCCGACTGGAGACCCAGCGCGCCTGTGCCCTGTCCAGGCCCTGCCCA	735
QY	963	GCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTA-AGGCCAACTGGGGATGCGGAT	1021
Db	736	CCCTCCAGGGGTGCGAGTCCACAAAACAGTGCCTTCTAGAGCCGGGCTGGGAATGGGGAC	795
QY	1022	ACAGGGCCTGCCATCCTCAGCAAAATGACCCATAGGACAGGCCCTGGACTGCTGGTAGATG	1081
Db	796	ACGGTGTCCACCATCCCCAGCTGGTGGCCCTGTGCCTGGGCCCTGGGCTGATGGAAGATG	855
QY	1082	CT-CTTCTCCATGCTCTGGCTGCAGTTAACTGTCTTCTTGGATTCACTGTGTAGAGC	1139
Db	856	GTCCGTGCCAGGCCCTTGGCTGCAGGCACACACTTAGCTTGGTTCACCATGAGCAAC	914

RESULT 4

US-09-182-145-14/C

03 03 102 143 14/C
; Sequence 14, Application US/09182145B

; sequence 14, Appl. No. 6387657; Patent No. 6387657

; GENERAL INFORMATION:

APPLICANT: Botstein, David A.

APPLICANT: Cohen, Robert

APPLICANT: Goddard, Audrey

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Lawrence, David A.

APPLICANT: Levine, Arnold J.

```

; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29.
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 14
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-14

```

Query Match	32.9%;	Score 561.4;	DB 4;	Length 1293;
Best Local Similarity	78.0%;	Pred. NO. 2.2e-152;		
Matches 701; Conservative	0;	Mismatches 196;	Indels 2;	Gaps 2;

QY	243	GGTGACATGAGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTCTCCCTCCCTGCTCTC	302
Db	1278	GGGACATGAGAGGCACACCGAAGACCCACCTCTGGCCCTTCTCCCTCCTCTGCTCCTC	1219
QY	303	TCAATGGTGTGCCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTGGACACCAACCCAG	362
Db	1218	TCAAAGGTGCTACCCAGCTGTGCCCGACACCCATGTACCTGCCCTGGCCACCTCCCCGA	1159
QY	363	TGCCACAGGGGTACCCCTGGTCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGG	422
Db	1158	TGCCCGCTGGAGTACCCCTGGTCTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG	1099
QY	423	AGGTGGGGAGTCTCGGACCACCTGCATGTCTCGGACCCCAAGCCAGGGCCCTGGTTGT	482
Db	1098	CGCTGGGGAGCTCGGACCAACTCCACGTCTCGGACGCCAGCCAGGGCCCTGGTCTGC	1039
QY	483	CAGCCTGGGGCAGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT	542
Db	1038	CAGCCCGGGCAGCACCCGGTGGCCGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGC	979
QY	543	AGCTGTGAGGTGATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG	602
Db	978	AGCTGTGAGGTGATCGGCCGCCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCCACTGCAGC	919
QY	603	GTCTGTGCCGCTGATGACGGTGGCTTACCTGCCCTGCCGTGTGCAGTGAAGATGTG	662
Db	918	ATCCGCTGCCGCTGGAGGACGGCGCTTACCTGCGTGCCTGTGCAGCGAGGATGTG	859
QY	663	CGGCTGCCCAGCTGGGACTGCCACGCCCCCAAGAAATACAGGTGCCAGGAAGTGTCTGC	722
Db	858	CGGCTGCCCAGCTGGGACTGCCCCCAACCCCAAGAGGGTCTGAGGTCTTGGCAAGTGTCTGC	799
QY	723	CCCGAGTGGGTATGTGACCAGGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGGA	782
Db	798	CCTGAGTGGGTGTGGGCCAAGGAGGGGACTGGGGACCCAGCCCCCTTCCAGCCCCAAGGA	739
QY	783	CACCAACTTCTGCGCTTGTCACTCCTGCCCTGCTGATGCTCCTTGTCCAAATTTGGAGC	842
Db	738	CCCCAGTTTCTGGGCTTGTCTCTTCCCTGCCCTGGTGTCCCTTCCCTGCCCCAGAAATGGAGC	679
QY	843	ACAGCCTGGGGCCCTGCTCAACCAACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAC	902
Db	678	ACGGCCTGGGGACCCCTGCTCGACCACTGTGGGCTGGGCATGGCCACCCCGGTGTCCAAC	619
QY	903	CAGAACCGATTCTGCAACTGGAGATCCAACGCCCGCCTGTGTCTGCCAGACCCCTGCCTG	962
Db	618	CAGAACCGCTTCTGCGGACTGGAGACCCAGCGCGGCTGTGCCTGTCCAGGCCCTGCCTCA	559
QY	963	GCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTTA-AGGCCAACTGGGGATGCGGAT	1021

QY 836 TTGGAGCACAGCCTGGGGCCCTGCTCAACCAACCTGTGGGCTGGGCATAGCCACCCGAGT 895
Db 732 CCAGACCACAGAGTGGAGCGCCTGTTCGAAGACCTGTGGGATGGGCATCTCCACCCGGGT 791
QY 896 GTCCAAACAGAACCGATTCTGCCAACTGGAGATCCAAACCGCGCTGTGTCTGCCCCAGACC 955
Db 792 TACCAATGACAACGCCCTCCTGCAGGCTAGAGAAGCAGAGCGCGCTGTGCATGGTCAAGGCC 851
QY 956 CTGCCTGGCAGCC 968
Db 852 TTGCGAAGCTGAC 864

RESULT 8

US-08-386-680-1
; Sequence 1, Application US/08386680
; Patent No. 5585270
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/386,680
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
US-08-386-680-1

Query Match 9.6%; Score 163.8; DB 1; Length 2075;
Best Local Similarity 53.5%; Pred. No. 2.1e-37;
Matches 392; Conservative 0; Mismatches 332; Indels 9; Gaps 2;
QY 242 CGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCTTCCTCTGCTTCT 301
Db 135 CGCGCCAGTATGGGCCCCGTCCGGGTGCGCCTTCGTGGTCTCTCGCCCTCTGCAGCCG 194
QY 302 CTCAATGGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCCA 361

Db 195 GCCGGCGCTCGGCCAGAACTGCAGCGGCGCGTCCCGTGCAGGAGCCGCGCGCG 254
QY 362 GTGCCACAGAGGGGTACCCCTGGTCTGGATGGCTGTGGCTGTGTAAAGTGTGTGACG 421
Db 255 CTGCCCGCGGGCGTGTGAGCCTGTGCTGGACGGCTGCGGCTGTGCGCGCTCTCGCCCAA 314
QY 422 GAGGCTGGGGAGTCTCTGCGACCACTGTCTGCGACCCCAAGCCAGGCGCTGTTT 481
Db 315 GCAGCTGGCGAGCTGTGCACCGAGCGGACCCCTGCGACCCGACAAAGGCGCTCTTCTG 374
QY 482 TCAGCCTGGGCGAGGCCCTGGCGGCATGGGGCTGTGTCTTGGATGAGGATGACGG 541
Db 375 TGACTTCGGCTCCCGGCCAACCGCAAGATCGCGTGTG---CACCGCCAAAGATGGTGC 431
QY 542 TAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG 601
Db 432 TCCCTGCATCTTCGGTGTACGGTGTACCGCAGCGGAGAGTCTCTCCAGAGCAGCTGCAA 491
QY 602 GGCTCTGTGCCGCTGTGATGACGGTGGCTTCACCTGCCCTGCGCTGTGCAGTGAGGATGT 661
Db 492 GTACCACTGACAGTGCCTGGACGGGCGGTGGCTGCATGCCCCCTGTGCAGCATGGACGT 551
QY 662 CGGCTGCCAGCTGGGACTGCCACACGCCCAAGAGAAATACAGTGCAGGAAAGTGCTG 721
Db 552 TCGTCTGCCAGCCCTGACTGCCCCCTTCCCGAGGAGGTCAAGTGCCTGGGAAATGCTG 611
QY 722 CCCCAGTGGGTATGTGACCAGGGAGTGACACCGGCGGATCCAGCGCTCCACGGCGCAAG 781
Db 612 CGAGGAGTGGGTGTGACGAGCCCAAGACCAACCGTGGTGGGCTGCCCTCGCGGC 671
QY 782 ACACCAACTTCTGCCCTTGTACTCCTCCTCTCTGC-----TGATGCTCCTTGTCCAAA 835
Db 672 TTACCGACTGGAAGACACGTTTGGCCCAAGACCCAACTATGATTAGAGCCAACTGCCTGGT 731
QY 836 TTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGT 895
Db 732 CCAGACCACAGAGTGGAGCGCCTGTTCGAAGACCTGTGGGATGGGCATCTCCACCCGGT 791
QY 896 GTCCAAACAGAACCGATTCTGCCAACTGGAGATCCAAACGCCGCTGTGTCTGCCCCAGACC 955
Db 792 TACCAATGACAACGCCCTCCTGCAGGCTAGAGAAGCAGAGCCCGCCTGTGCATGGTCAAGGCC 851
QY 956 CTGCCTGGCAGCC 968
Db 852 TTGCGAAGCTGAC 864

RESULT 9

US-08-459-717-1
; Sequence 1, Application US/08459717
; Patent No. 5770209
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,717
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,427
; FILING DATE: 30-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
;
US-08-459-717-1

```

		Query Match										
		9.6%; Score 163.8; DB 1; Length 2075;										
		Best Local Similarity 53.5%; Pred. No. 2.1e-37;										
		Matches 392; Conservative 0; Mismatches 332; Indels 9; Gaps 2;										
QY	242	CGGTGACATGAGGGG	CAGCCC	CACTGATCCATCTCTGGCC	CACTTCCTT	CCTCTG	CCTTCT	301				
Db	135	CGCGCCAGTATGGG	CCCCCG	TCGCGCTCGCT	CGCTTCTG	TGGTCTCT	CGCCCTCT	194				
QY	302	CTCAATGGTGTGTG	CCCCAGCTGT	CGCGGACAC	CCCTGTAC	CTGTCTCT	TGGACAC	361				
Db	195	GCCGGCCGTGCG	CCAGAACTG	CAGCGGGCCGT	GCCGGTCC	CGGACGAG	CGCGCGCG	254				
QY	362	GTGCCCACAGGGG	TACCCCTGGT	GCTGGATGG	CTGTGGCT	GCTGTAA	AGTGTGTG	421				
Db	255	CTGCCCGGCGSG	CGTGAGCCTCG	TGCTGGAC	GGCTGCGGCT	GCTGCC	CGCTCTG	314				
QY	422	GAGGCTGGGGAG	TCCCTGCG	ACCACCTGC	ATGTCTCG	CAACCCCA	GCCAGGG	481				
Db	315	GCAGCTGGGCGA	GCTGTGC	ACCGAGCG	CAACCCCTG	CGACCCG	CAACAAGG	374				
QY	482	TCAGCCTGGGCA	GCCCCCTGG	CGGGCCATGG	GGGCTGTGT	GTCTCTT	TGGATG	541				
Db	375	TGACTTCGGCT	CCCCGGCCAA	CCGCAAGAT	CGGGCTGTG	---CAC	CGCCAAAG	431				
QY	542	TAGCTGTGAGGT	GAATGGCCG	CAGGTACCTGG	ATGGAGAG	ACCTTTAA	ACCCAAT	601				
Db	432	TCCTTGCACT	TCGGTGGTAC	GGTGATCC	CGAGCGAG	AGATCTCT	CCAGAGCA	491				
QY	602	GGTCTGTGCCCT	GTGATGAC	CGGTGGCTTCA	CTGCTGCCT	GTGCTGTG	CAGTGAGG	661				
Db	492	GTACCAGTGC	ACGTGCC	TGGACGGGG	CGGTGGGCTGC	ATGCCCCCTGT	GCAGCATG	551				
QY	662	GCGGCTGCCAG	TGGGACTGCC	CCACGCCCA	AGAGAA	TACAGGTGCC	CAGGAA	721				
Db	552	TCGTCTGCCAG	CCCTGACTG	CCCCCTTCC	CGAGGAGG	GTCAAGCTG	CCCCGGAA	611				
QY	722	CCCCGAGTGGG	TATGTGAC	CAGGGAGTGAC	ACCGGCGAT	CCAGCGCTCC	ACGGCGCA	781				
Db	612	CGAGGAGTGGG	TGTGAC	GAGCCCCA	AGGACCA	ACCGTGGT	TGGCCCTG	671				
				</								

Db		792	TACCAATGACAAGCCTCCTGCAGGCTAGAGAAGCAGCGCCCTGTGCATGGTCAGGCC	851
QY		956	CTGCCTGGCAGCC	968
Db		852	TTGCGAAGCTGAC	864

RESULT 10
US-08-712-302-1
; Sequence 1, Application US/08712302
; Patent No. 5783187
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,302
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,680
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
US-08-712-302-1

	Query Match	9.6%;	Score 163.8;	DB 1;	Length 2075;
	Best Local Similarity	53.5%;	Pred. No. 2.1e-37;		
	Matches 392; Conservative	0;	Mismatches 332;	Indels	Gaps 2;
QY	242	C GGTGACATGAGGGCAGCCCACTGATCCAATCTTCTGGCCACTTCCTTCCCTCTGCCTTCT	301		
Dδ	135	C GCCGCCAGTAGGGCCCCCGTGCCGGCTCGCTGGTGTCTCCTCGCCCCTCTGCAGCCG	194		
QY	302	CTCAATGGTGTGTGCCCAAGCTGTGCCGGACACCCTGTACCTGTCTTGGACACCACCCCA	361		
Dδ	195	GCCGGCCGTGGCGAAGAATGCAGCGGGCCGTGCCGGTGCCTGGACGAGCCGGCGCCGCG	254		

QY 956 CTGCCTGGCAGCC 968
Db 852 TTGCGAAGCTGAC 864

RESULT 12
US-09-097-179-1
; Sequence 1, Application US/09097179
; Patent No. 6149916
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,179
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,680
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
US-09-097-179-1

Query Match 9.6%; Score 163.8; DB 3; Length 2075;
Best Local Similarity 53.5%; Pred. No. 2.1e-37;
Matches 392; Conservative 0; Mismatches 332; Indels 9; Gaps 2;

QY 242 CGGTGACATGAGGGCAGCCCACTGATCCATCTTCTTGCCCACTTCCTTCCTCTGCTTCT 301
Db 135 CGCCGCCAGTATGGGCCCCCGTCCCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 194

QY 302 CTCAATGGTGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCA 361
Db 195 GCCGGCCGTGCGCAGAACTGCAGCGGGCCGCTGCCGGTGCCCGGACGAGCCGGCGCGG 254

QY 362 GTGCCCCACAGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGTGTAAGTGTGTGCACG 421
Db 255 CTGCCCGGGCGGCTGAGCCTCGTGTGGACGGCTGCGGCTGCTGCCCGCTCTGCCGCCAA 314

QY 422 GAGGCTGGGGAGTCTCTGCGACCACCTGTCATGTCTGCGACCCCGAGCCAGGGCCTGTTTG 481
Db 315 GCAGCTGGGGAGTGTGACCGGCGGACCCCTCGACCCCGCACAGGGCCTCTTCTG 374

QY 482 TCAGCCTGGGCGAGCCCTGGCGGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGG 541
Db 375 TGACTTCGGCTGCCCCGCCAACCGCAAGATCGGGCTGTG---CACCGCCAAAGATGGTGC 431

QY 542 TAGCTGTGAGCTGAATGGCCGCGAGGTACTTGGATGGAGAGACCTTTAAACCCAATTCAG 601
Db 432 TCCCTGCATCTTCGGTGGTACGGTGTACCGAGCGGAGAGTCTTCCAGAGCAGCTGCAA 491

QY 602 GGTCTGTGCCGTGTGATGACGGTGGCTTACCTGCCTGCCGTGCGAGTGAGGATGT 661
Db 492 GTACCAGTGCAGTGCCTGGACGGGGCGGTGGGCTGCATGCCCTGTGAGCATGGACGT 551

QY 662 GCGGCTGCCAGCTGGGACTGCCACGCGCCCAAGAGAAATACAGGTGCCAGGAAAGTGTG 721
Db 552 TCGTCTGCCAGCCCTGACTGCCCTTCCCGAGGAGGGTCAAGCTGCCGGGAAATGCTG 611

QY 722 CCCCAGTGGGTATGTGACCAGGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGG 781
Db 612 CGAGGAGTGGGTGTGACGAGCCCAAGGACCAACCGTGGTGGGCTTGGCCTCGCCCTCGCGG 671

QY 782 ACACCAACTTCTGCCCCCTGTGCTACTCCTGCCTCTGC-----TGATGCTCTTGTCCAA 835
Db 672 TTACCGACTGGAGACACAGTGTGGCCCCAGACCCCAACTATGATTAGAGCCCAACTGCCTGGT 731

QY 836 TTGGAGCACAGCTGGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCTAGCCACCCGAGT 895
Db 732 CCAGACCACAGAGTGGAGCGCCTGTTCACAGACCTGTGGGATGGGCATCTCCACCCGGGT 791

QY 896 GTCCAACCAGAAAGCATTTCTGCCAACTGGCAATCCAACGCCGCTGTGTGTGCCCCAGACC 955
Db 792 TACCAATGACAAAGCCCTCCTGCAGGCTAGAGAAAGCAGAGCCGCTGTGATGGTCAAGGCC 851

QY 956 CTGCCTGGCAGCC 968
Db 852 TTGCGAAGCTGAC 864

RESULT 13
US-09-080-715-1
; Sequence 1, Application US/09080715
; Patent No. 6190884
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,715
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:36:41 ; Search time 421.038 Seconds
(without alignments)
8368.853 Million cell updates/sec

Title: US-10-010-408-1
Perfect score: 1708
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GCCTAGATAAACACCCAAA 1708

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1708	100.0	1708	15	US-10-010-408-1 Sequence 1, Appli
2	1278	74.8	1734	15	US-10-112-267-17 Sequence 17, Appl
3	1278	74.8	1734	15	US-10-112-267-18 Sequence 18, Appl
4	753	44.1	753	15	US-10-010-408-3 Sequence 3, Appli
5	681	39.9	681	15	US-10-010-408-12 Sequence 12, Appl
6	566.8	33.2	1337	11	US-09-915-582-30 Sequence 30, Appl
7	566.6	33.2	1352	11	US-09-915-582-14 Sequence 14, Appl
8	561.4	32.9	1266	14	US-10-137-866-319 Sequence 319, App
9	561.4	32.9	1266	14	US-10-146-726-319 Sequence 319, App
10	561.4	32.9	1266	14	US-10-146-727-319 Sequence 319, App
11	561.4	32.9	1266	14	US-10-146-788-319 Sequence 319, App
12	561.4	32.9	1266	14	US-10-152-380-319 Sequence 319, App
13	561.4	32.9	1266	14	US-10-153-934-319 Sequence 319, App
14	561.4	32.9	1266	15	US-10-028-072-319 Sequence 319, App
15	561.4	32.9	1266	15	US-10-121-049-319 Sequence 319, App
16	561.4	32.9	1266	15	US-10-123-904-319 Sequence 319, App

17	561.4	32.9	1266	15	US-10-140-470-319	Sequence 319, App
18	561.4	32.9	1266	15	US-10-175-746-319	Sequence 319, App
19	561.4	32.9	1266	15	US-10-176-918-319	Sequence 319, App
20	561.4	32.9	1266	15	US-10-176-921-319	Sequence 319, App
21	561.4	32.9	1266	15	US-10-137-865-319	Sequence 319, App
22	561.4	32.9	1266	15	US-10-140-474-319	Sequence 319, App
23	561.4	32.9	1266	15	US-10-142-431-319	Sequence 319, App
24	561.4	32.9	1266	15	US-10-143-114-319	Sequence 319, App
25	561.4	32.9	1266	15	US-10-140-002-319	Sequence 319, App
26	561.4	32.9	1266	15	US-10-142-419-319	Sequence 319, App
27	561.4	32.9	1266	15	US-10-123-262-319	Sequence 319, App
28	561.4	32.9	1266	15	US-10-142-423-319	Sequence 319, App
29	561.4	32.9	1266	15	US-10-121-050-319	Sequence 319, App
30	561.4	32.9	1266	15	US-10-141-755-319	Sequence 319, App
31	561.4	32.9	1266	15	US-10-143-032-319	Sequence 319, App
32	561.4	32.9	1266	15	US-10-123-108-319	Sequence 319, App
33	561.4	32.9	1266	15	US-10-123-236-319	Sequence 319, App
34	561.4	32.9	1266	15	US-10-123-261-319	Sequence 319, App
35	561.4	32.9	1266	15	US-10-140-921-319	Sequence 319, App
36	561.4	32.9	1266	15	US-10-140-928-319	Sequence 319, App
37	561.4	32.9	1266	15	US-10-121-045-319	Sequence 319, App
38	561.4	32.9	1266	15	US-10-123-292-319	Sequence 319, App
39	561.4	32.9	1266	15	US-10-123-903-319	Sequence 319, App
40	561.4	32.9	1266	15	US-10-124-819-319	Sequence 319, App
41	561.4	32.9	1266	15	US-10-124-822-319	Sequence 319, App
42	561.4	32.9	1266	15	US-10-140-925-319	Sequence 319, App
43	561.4	32.9	1266	15	US-10-160-498-319	Sequence 319, App
44	561.4	32.9	1266	15	US-10-124-824-319	Sequence 319, App
45	561.4	32.9	1266	15	US-10-127-825A-319	Sequence 319, App

ALIGNMENTS

RESULT 1

US-10-010-408-1
; Sequence 1, Application US/10010408
; Publication No. US20020165185A1

GENERAL INFORMATION:

APPLICANT: John J. Castellet, Jr.
TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules and Uses Therefor

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/010,408

FILING DATE: 07-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/044,273

FILING DATE: March 19, 1998

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: MBI-004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

; LENGTH: 1708 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 249..1001
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-010-408-1

Query Match		100.0%;	Score 1708;	DB 15;	Length 1708;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1708;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GACGCTTCTGATCTCCAGAGGACCCCTGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCC	60		
Db	1	GACGCTTCTGATCTCCAGAGGACCCCTGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCC	60		
QY	61	GCTGGGCAGTGGCTTGAATGGAGGTCTTATTAATCTGGGAACCTGAGGAGCTAAGAGGCTC	120		
Db	61	GCTGGGCAGTGGCTTGAATGGAGGTCTTATTAATCTGGGAACCTGAGGAGCTAAGAGGCTC	120		
QY	121	CTGTGAGCTTGCTCTAAAGTCTTAGCACCTTGCTGGTGGCTTGGGCTTCACACACTGTCAGA	180		
Db	121	CTGTGAGCTTGCTCTAAAGTCTTAGCACCTTGCTGGTGGCTTGGGCTTCACACACTGTCAGA	180		
QY	181	CACCTTCGTGGTGGCTCCACGGCCTCACCTTCAGGTTTGAAGCTGGCTCCACACAGGGAC	240		
Db	181	CACCTTCGTGGTGGCTCCACGGCCTCACCTTCAGGTTTGAAGCTGGCTCCACACAGGGAC	240		
QY	241	ACGGTGACATGAGGGGCAGCCCACTGATCCATCTCTGGCCACTTCCTCTGCTCTGCTTC	300		
Db	241	ACGGTGACATGAGGGGCAGCCCACTGATCCATCTCTGGCCACTTCCTCTGCTCTGCTTC	300		
QY	301	TCTCAATGGTGTGCCCCAGCTGTGCCGACACCTGTACCTGTCTGGCCACTTCCTCTGCTTC	360		
Db	301	TCTCAATGGTGTGCCCCAGCTGTGCCGACACCTGTACCTGTCTGGCCACTTCCTCTGCTTC	360		
QY	361	AGTGCCCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGCTGCTGCTGCTGCTG	420		
Db	361	AGTGCCCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGCTGCTGCTGCTGCTG	420		
QY	421	GGAGGCTGGGGAGTCTCGACACACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480		
Db	421	GGAGGCTGGGGAGTCTCGACACACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480		
QY	481	GTCAGCCTGGGCGAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACG	540		
Db	481	GTCAGCCTGGGCGAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACG	540		
QY	541	GTAGCTGTGAGTGAATGGCCGACCTGCTGATGGAGAGACCTTTAAACCCCAATTGCA	600		
Db	541	GTAGCTGTGAGTGAATGGCCGACCTGCTGATGGAGAGACCTTTAAACCCCAATTGCA	600		
QY	601	GGGTCTCTGCGGCTGTGATGACGGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG	660		
Db	601	GGGTCTCTGCGGCTGTGATGACGGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG	660		
QY	661	TGCGGCTGCCAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGGCAGGAAAGTGCT	720		
Db	661	TGCGGCTGCCAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGGCAGGAAAGTGCT	720		
QY	721	GCCCCGAGTGGGTATGTACACAGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAG	780		
Db	721	GCCCCGAGTGGGTATGTACACAGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAG	780		
QY	781	GACACCAACTTCTGCCCCCTGTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	840		
Db	781	GACACCAACTTCTGCCCCCTGTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	840		
QY	841	GCACAGCCTGGGGCCCCCTGCTCAACCCACTGTGGGCTGGGCATAGCCACCCGAGTGCCA	900		

Db	841	GCACAGCCTGGGGCCCCCTGCTCAACCCACTGTGGGCTGGGCATAGCCACCCGAGTGCCA	900
QY	901	ACCAGAACCGATTCTGCCAACTGGAGATCCAAACGCCGCCCTGTGTCTGCCACAGACCCTGCC	960
Db	901	ACCAGAACCGATTCTGCCAACTGGAGATCCAAACGCCGCCCTGTGTCTGCCACAGACCCTGCC	960
QY	961	TGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCAACTGGGGATGCGGA	1020
Db	961	TGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCAACTGGGGATGCGGA	1020
QY	1021	TACAGGGCCCTGCATCCTCAGCAAAATGACCTTAGGACGAGCCCTGGACTGCTGGTAGAT	1080
Db	1021	TACAGGGCCCTGCATCCTCAGCAAAATGACCTTAGGACGAGCCCTGGACTGCTGGTAGAT	1080
QY	1081	GCTCTTCTCCATGCTCTTGGCTGCAGTTAACTGTCTCTGCTTGGATTCACTGTGTAGAGCC	1140
Db	1081	GCTCTTCTCCATGCTCTTGGCTGCAGTTAACTGTCTCTGCTTGGATTCACTGTGTAGAGCC	1140
QY	1141	ACTGAGCGGATCCTGCTCTGTCTGAGGTAGGCGGAGCAGGTGACCACTCCAGTTCTCTG	1200
Db	1141	ACTGAGCGGATCCTGCTCTGTCTGAGGTAGGCGGAGCAGGTGACCACTCCAGTTCTCTG	1200
QY	1201	GTTGAGCCTGGAAATCTTGGGTTCTCTGGCTCATTCCTCAAAACATCCCTGTACAAAAAG	1260
Db	1201	GTTGAGCCTGGAAATCTTGGGTTCTCTGGCTCATTCCTCAAAACATCCCTGTACAAAAAG	1260
QY	1261	GACAACCAAAAGACCTTTAAACCTTAGGACGACAGCAGATTCCTGAAACTTCCAAATCCCTTGG	1320
Db	1261	GACAACCAAAAGACCTTTAAACCTTAGGACGACAGCAGATTCCTGAAACTTCCAAATCCCTTGG	1320
QY	1321	GATAAGGTCAATCTTAGGACGACAGCAGATTCCTGAAACTTCCAAATCCCTTGG	1380
Db	1321	GATAAGGTCAATCTTAGGACGACAGCAGATTCCTGAAACTTCCAAATCCCTTGG	1380
QY	1381	ACTTCTGTATGCTTGTCTCCCAAGATGATGAATGAACTCGTAAGTGTACCTTCCCTGACC	1440
Db	1381	ACTTCTGTATGCTTGTCTCCCAAGATGATGAATGAACTCGTAAGTGTACCTTCCCTGACC	1440
QY	1441	TGAGAACACCCCTGCTGCTCGGGAAAGTATTCAGGGGCGAGAATTCCTGTGAACATGAAGA	1500
Db	1441	TGAGAACACCCCTGCTGCTCGGGAAAGTATTCAGGGGCGAGAATTCCTGTGAACATGAAGA	1500
QY	1501	GATGAATCACACTGTCTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTTGTATTTT	1560
Db	1501	GATGAATCACACTGTCTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTTGTATTTT	1560
QY	1561	CAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAGGCTCCACACCTCTGGCAGGCC	1620
Db	1561	CAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAGGCTCCACACCTCTGGCAGGCC	1620
QY	1621	AGGGCCTTCTCTTCAGCATGAGAAAGACAAGGGACAGCAGAGTACTCTCTCTGGAGGA	1680
Db	1621	AGGGCCTTCTCTTCAGCATGAGAAAGACAAGGGACAGCAGAGTACTCTCTCTGGAGGA	1680
QY	1681	CTAGTCTAGCCTAGAATAAACACCCCAA 1708	
Db	1681	CTAGTCTAGCCTAGAATAAACACCCCAA 1708	

RESULT 2
US-10-112-267-17
; Sequence 17, Application US/10112267
; Publication No. US2003006878A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann

; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 18
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-112-267-18

Query Match 74.8%; Score 1278; DB 15; Length 1734;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 1520; Conservative 0; Mismatches 165; Indels 33; Gaps 11;

QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGC 62
Db 1722 CGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCCCTTGGCGAGGCTGCAGCTGC 1663

QY 63 TG-GGCAGTGGCTTGGAAATGGAGGTCTTTATTAATGGAACCTGAGGAGCTAAAGAGGCTCC 121
Db 1662 TGTGGCAGTAGCTGGGATGGAGGTCTTTCTTGTCTGGGAACCTGAGAGGCTGCAGGCTCC 1603

QY 122 TGTGAG---CTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCA 178
Db 1602 TGTGAGGCTCCTGTCTCTAAACTCTTGGCACTTGGGCTGGGCTTGGGCTTCACACACTGTCA 1543

QY 179 GACACCTTCGTGGTGGCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238
Db 1542 GACACCTTCGTGGTGGCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 1489

QY 239 ACACGGTGACATGAGGGGCGACGCCCACTGATCCATCTTCTGGCCACTTCCTCTGCGCT 298
Db 1488 ACACGGTGACATGAGGGGCAACCCCACTGATCCATCTTCTGGCCATTCCTCTCTGCGAT 1429

QY 299 TCTCTCAATGGTGTGTCGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCAACC 358
Db 1428 TCTCTCAATGGTGTATTCACAGCTGTGCCAGCACCCCTGTGCCCTGTCTTGGACACCAACC 1369

QY 359 CCAGTGCCACAGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTCTGTAAGTGTGTGC 418
Db 1368 CCAGTGCCACCGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTCTGCGAGTGTGTGC 1309

QY 419 ACGGAGGCTGGGGAGTCTGCGACCACTGTCATGTCTGCGACCCCAAGCCAGGGCCTGGT 478
Db 1308 ACGGAGGCTGGGGAGTCTGCGACCACTGTCATGTCTGCGACCCCAAGCCAGGGCCTGGT 1249

QY 479 TTGTCAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGA 538
Db 1248 TTGTCAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGA 1189

QY 539 CGGTAGCTGTAGGTGAATGGCGCGAGGTACCTGGATGGAGAGACCTTTAAACCCAATTG 598
Db 1188 CGGGAGCTGTAGGTGAATGGCGCGAGGTACCTGGATGGGAGACCTTTAAACCCAATTG 1129

599 CAGGGTCTCTGTCGGCTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGCAGTGAGGA 658
Db 1128 CAGGGTCTCTGTCGGCTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGCAGTGAGGA 1069

QY 659 TGTGCGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTG 718
Db 1068 TGTGCGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTG 1009

QY 719 CTGCCCCGAGTGGGTATGTGACCAAGGAGTGA---CACCGCGGATCCAGCGCTCCACGGC 775
Db 1008 CTGCCCCGAGTGGGTATGTGACCAAGGAGTGAATGACCGCGCAATCCAGCCCTCCTCAGC 949

QY 776 GCAAGGACACCAACTTTCTGCCCTTGTCACTCTCTGCCTCTCTGCTGATGCTCTTGTCCAAA 835
Db 948 CCAAGGACACCAACTTTCTGCCCTTGTCACTCTCTGCATCTGCCGATGCCCTGTCCAAA 889

QY 836 TTGGAGCACAGGCTGGGGCCCTGCTCAACCACTGTGGGCTGGSCATAGCCACCCGAGT 895
Db 888 CTGGAGCACAGGCTGGGGCCCTGCTCAACCACTGTGGGTTGGSCATAGCCACCCGAGT 829

QY 896 GTCCAAACCAGAACCGATTCTGCCAACTGGAGATCCAACGCCGCTGTGTCTGCCCCAGACC 955
Db 828 ATCCAAACCAGAACCGATTCTGCCAACTGGAGATCCAACGCCGCTGTGTCTGTCTCCAGACC 769

QY 956 CTGCTGCGAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTTTAAGGCCA-ACTGGGGA 1014
Db 768 CTGCTGCGATCCAGGAGCCACGGCTCATGGAACAGTGTCTTTCTAGAGCCATTGCGGGGA 709

QY 1015 TCGGATACAGGSCCTGCCATCTCAGCAAAATGACCCCTAGGACCAAGGCCCTGGACTGCTG 1074
Db 708 TGTGGATACAGGSCCTGCCATCTCAGCAAAATGTCCCTAGGACCAAGGCCCTGGACTGATG 649

QY 1075 GTAGATGCTCTTGTCCATGCTCTTGGCTGCAGTTAACTGTCTCTTGGATTCACTGTGT 1134
Db 648 GTAGATGCTCTTGTCCATGCTCTTGGCTGCAGTTAACTGTCTCTTGGATTCACTGTGT 589

QY 1135 AGAGCCACTGAGGATCCCTGCTCTGTCTGAGGTAGGGCGGAGCAGGTGACCACTCCAGT 1194
Db 588 AGAGCCCTGAGGATCCCTGCTCTGTCTGAGGTGGGGGAGCAGGTGACCACTCCAGT 529

QY 1195 TCTCTGG--TTCAACCTGGAAATTCCTGGGTTCTCTGGCTCATTCCTCAAAACATCCCTGT 1252
Db 528 TCTCTGGATTCTGACCCAGGCTTCTGGGTTCTCTGGCTAGTTCTCTCAAAACTTCCCTGT 469

QY 1253 ACAAAAGGACAAACCAAAAGACCTTTAAACCTAGGCTATACCTGGCAAAACCTGGCCACC 1312
Db 468 ATGAAAAGGACAAACCAAAAGGACCTTTAAAGCTAAGCTGTACTGGCAAGCCTTGGCCACC 409

QY 1313 GTGCTGGGATAAGGTCAATGTTAGG-ACCAGACAGCAGATTGCTCTGAAACTTCCAATTTC 1371
Db 408 ATGCTGGGATAGTGACAGTAATAGTACCAGGCAGCAGATTGCTCTGAAACATCCAGGTC 349

QY 1372 CCTTCTTGACTGTGTATGCTTGTCCCAAAAGATGATGAATGAATCGTAAGTGTACCT 1431
Db 348 CCTTCTTGACTGTGTATGCTGTGT-CCCAAGATTATGGGTGACCTGTGAAGTGTGCT 290

QY 1432 TCCCTGACCTGAGNACACCCCTGCCCTGCTCGGGGAAGTATTTCAGSSGSCAGAAATTCCTGTGA 1491
Db 289 TTCTCTGATCTGAGNACACCCCTGCCCGGCTGGGA-----AGAATTTCTGGGA 243

QY 1492 ACATGAAGAGAT-GAATCACACTGTCTCTTAAGAAATTCCTGAAAGTCCAGGAACTTGAGC 1550
Db 242 ACATGAAGAGATGGAATCACACTATTCTTAAGAGCGTTTGGCAAGTCCAGGAACTTGACC 183

QY 1551 TTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCT 1610
Db 182 TTTGTATTTGTAATAATACACATCTCTTAATGTCTCACAAAGCAAG-AGGCTCCACACTT 124

QY 1611 CTGGCAGGCCAGGGGCTTTCTCTCTTCTCAGCATGAGAAAGCAAGGACAGCAGACTCTC 1670
Db 123 CTGGCAGGCCAGGGGCTTTCTCTCTTCTCAGCATGAGAGACAAAGNACAGTAGAGTACCCTC 64

QY 1671 CTCTGGAGGACTAGTCTAGCCTAGAAATAAACACCCCAA 1708


```
;
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 681 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..681
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-010-408-12

Query Match      39.9%; Score 681; DB 15; Length 681;
Best Local Similarity 100.0%; Pred. NO. 1.1e-202;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 CAGCTGTGCCGGACACCCTGTACCTGTCTTGGACACACCACCCAGTGCCACAGGGGTA 377
Db 1 CAGCTGTGCCGGACACCCTGTACCTGTCTTGGACACACCACCCAGTGCCACAGGGGTA 60

QY 378 CCCCTGGTGTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGSGTGGGGAGTCC 437
Db 61 CCCCTGGTGTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGSGTGGGGAGTCC 120

QY 438 TGCACACCCTGCATGTCTGCACCCCCAGCCAGGGCCCTGGTTTGTCAACCTGGGGCAGGC 497
Db 121 TGCACACCCTGCATGTCTGCACCCCCAGCCAGGGCCCTGGTTTGTCAACCTGGGGCAGGC 180

QY 498 CCTGGCGGCCATGGGGCTGTGTGTCTCTTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT 557
Db 181 CCTGGCGGCCATGGGGCTGTGTGTCTCTTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT 240

QY 558 GGCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTGCAGGGTCTGTGCCGCTGT 617
Db 241 GGCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTGCAGGGTCTGTGCCGCTGT 300

QY 618 GATGACGGTGGCTTACCTGCCTGCCCTGTGCAGTGGAGATGTGCGGTGCCAGCTGG 677
Db 301 GATGACGGTGGCTTACCTGCCTGCCCTGTGCAGTGGAGATGTGCGGTGCCAGCTGG 360

QY 678 GACTGCCACGCCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGCCCCGAGTGGGTATGT 737
Db 361 GACTGCCACGCCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGCCCCGAGTGGGTATGT 420

QY 738 GACCAGGGAGTGACACCGGGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCC 797
Db 421 GACCAGGGAGTGACACCGGGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCC 480

QY 798 CTTGTCACTCCTGCTCTGCTGATGCTCCTGTGTCCAAATTGGAGCACAGCCTGGGGCCCC 857
Db 481 CTTGTCACTCCTGCTCTGCTGATGCTCCTGTGTCCAAATTGGAGCACAGCCTGGGGCCCC 540

QY 858 TGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAACCGATTCTGC 917
Db 541 TGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAACCGATTCTGC 600

QY 918 CAACTGGAGATCCAACGCGCCTGTGTCTGCCAGACCCCTGCCTGGCAGCCAGGAGCCAC 977
Db 601 CAACTGGAGATCCAACGCGCCTGTGTCTGCCAGACCCCTGCCTGGCAGCCAGGAGCCAC 660

QY 978 AGCTCATGGAACAGTCTTTC 998
Db 661 AGCTCATGGAACAGTCTTTC 681
```

RESULT 6

```
US-09-915-582-30
; Sequence 30, Application US/09915582
; Patent No. US20020120103A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1337
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1337)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-915-582-30

Query Match      33.2%; Score 566.8; DB 11; Length 1337;
Best Local Similarity 77.0%; Pred. NO. 8.4e-167;
Matches 728; Conservative 3; Mismatches 205; Indels 10; Gaps 3;

QY 196 CTCACGGCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGGACACGGTGACATGAGGG 255
Db 2 CTTACAGTTTCACCTTCAGGCTCAAARCTGGSTCTGCA-----GGGGACATGAGAG 53

QY 256 GCAGCCCCACTGATCCATCTTCTGGCCACTTCCTTCTCCTCTCTCAATGGTGTGTG 315
Db 54 GCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTCTCAAGGTGCCTA 113

QY 316 CCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCACCCAGTGCCACAGGGGG 375
Db 114 CCCAGCTGTGCCGGACACCCATGTACCTGCCCCCTGCCACCTCCCCGATGCCGCTGGGAG 173

QY 376 TACCCCTGGTCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGAGGCTGGGGAGT 435
Db 174 TACCCCTGGTCTGGATGGCTGTGGCTGTGCCGGTATGTGCACGCGGGCTGGGGGAGC 233

QY 436 CCTGCGACCACTGCATGTCTGCGACCCCGCAGCCAGGGCCCTGGTTTGTACGCCCTGGGGCAG 495
Db 234 CCTGCGACCAACTCCACGTCTGCGACGCGCAGCCAGGGCCCTGGTCTGCCAGCCCCGGGCGAG 293

QY 496 GCCCTGGCGGCCATGGGGCTGTGTGTCTTGGATGAGGATGACGGTAGCTGTGAGGTGA 555
Db 294 GACCCGGTGMCGGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGAGCTGTGAGGTGA 353

QY 556 ATGGCCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTGTGCCGCT 615
Db 354 ACGGCCGCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCT 413

QY 616 GTGATGACGGTGGCTTACCTGCCTGCCGCTGTGCAGTGAGGATGTGCGGCTGCCAGCT 675
Db 414 GCGAGGACGCGGCTTACCTGCCTGCCGCTGTGCAGGAGGATGTGCGGCTGCCAGCT 473

QY 676 GGGACTGCCACGCCCCCAAGAGAATACAGTGCCAGGAAAGTGTGCCCCGAGTGGGTAT 735
Db 474 GGGACTGCCCCACCCAGGAGGCTCGAGGTCCTGGGCAAGTGTGCCCCCTGAGTGGGTGT 533

QY 736 GTGACCAAGGAGTGACACCGGGGATCCAGCGCTCCACGGGCGCAGGACACCAACTTTCTG 795
Db 534 GCGGCCAAGGAGGGGGACTGGGGACCCAGCCCCCTTCCAGCCCCAAGGACCCAGTTTCTG 593

QY 796 CCCTTGCTACTCCTCTCCTCTGTGTGATGCTCTTGTCCAAATTGGAGCACAGCCTGGGGCC 855
Db 594 GCCTTGCTCTCTCCCTGCCCCCTGGTGTCTCCCTGCCCAGATGGAGACAGGCTGGGGAC 653
```

QY 856 CCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAACCGATTCT. 915
Db 654 CCTGCTCGACCACCTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAGAACCGCTTCT 713
QY 916 GCCAACTGAGATCCAACGCCGCTGTGTCTGCCAGACCCCTGCCTGGCAGCCAGGAGCC 975
Db 714 GCGACTGAGACCCAGCGCCGCTGTGCCTGTCCAGGCCCTGCCACCCCTCCAGGGGTC 773
QY 976 ACAGCTCATGGAACAGTGTCTTTCTA-AGGCCAACTGGGGATGCGGATACAGGGCCTGCCA 1034
Db 774 GCAGTCCACAAAACAGTGCCTTCTAGAGCCGGCTGGGAATGGGACACAGGTGTCCACCA 833
QY 1035 TCCTCAGCAAAATGACCCCTAGGACCAAGGCCCTGGACTGTGCTGATGCT-CTTCTCCATG 1093
Db 834 TCCCCAGCTGGTGGCCCTGTGCCTGGGCCCTGGGCTGATGGAAGATGGTCCGTGCCCCAGG 893
QY 1094 CTCTTGGCTGCAGTTAACTGTCTCTGCTTGGATTCACTGTGTAGAGC 1139
Db 894 CCCTTGGCTGCAGGCAACACTTTAGCTTGGGTCCACCATGCAGAAC 939

RESULT 7

US-09-915-582-14
; Sequence 14, Application US/09915582
; Patent No. US20020120103A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1352
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-915-582-14

Query Match 33.2%; Score 566.6; DB 11; Length 1352;
Best Local Similarity 77.0%; Pred. No. 9.7e-167;
Matches 732; Conservative 0; Mismatches 209; Indels 10; Gaps 3;
QY 191 GTGGCTCCACGGCCTCACCTTCAAGTTTGAAGCTGGCTCCACAGGGACACGGTGACAT 250
Db 3 GTCCGCTTCACAGTTTACCTTCAAGCTCAAAGCTGGCTCTGCA-----GGGGACAT 54
QY 251 GAGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTGCTCTCTCAATGGT 310
Db 55 GAGAGGCACACCGAAGACCCACCTCTGGCCTTCTCCCTCTCTGCTGCTCTCTCAAGGT 114
QY 311 GTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTCCCCACA 370
Db 115 GCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCCCTGGCCACCTCCCCGATGCCCGCT 174
QY 371 GGGGTACCCCTGGTGTGGATGCTGTGGCTGTAAAGTGTGTGCACGGAGGCTGGG 430
Db 175 GGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTGCCGGGTATGTGCACGGCGGCTGGG 234
QY 431 GGAGTCTCGACCACTGCATGTCTGCGACCCCAAGCCAGGCGCTGGTTGTGACGCTGG 490
Db 235 GGAGCCTGCGACCAACTCCACGCTGCGACGCCAGCCAGGCGCTGGTCTGCCAGCCCGG 294
QY 491 GGCAGGCCCTGGCGCCATGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGA 550

Db 295 GGAGGACCCGGTGGACGGGGGGCCCTGTGCTCTTTGGCAGAGGACGACAGCAGCTGTGA 354
QY 551 GGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATTCAGAGGTCTCTGTG 610
Db 355 GGTGAACGGCCCGCCTGTATCGGAAGGGGAGACCTTCCAGCCCACTGCAGCATCCGCTG 414
QY 611 CCGCTGTGATGACGGTGGCTTCACTGCTGCCCTGCCGCTGTGCAGTGAAGTGTGCGGCTGCC 670
Db 415 CCGCTGCGAGGACGGCGGCTTCACTGCTGCCGCTGCCGCTGTGCAGCAGGATGTGCGGCTGCC 474
QY 671 CAGCTGGGACTGCCACAGCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGCCCCGAGTG 730
Db 475 CAGCTGGGACTGCCCCCAACCCAGGAGGGTTCGAGGTCTTGGCAAGTGTGCCCCGAGTG 534
QY 731 GGTATGTGACCAAGGAGTGACACCGCGCGATCCAGCGCTCCACGGCGCAAGGACACCAACT 790
Db 535 GGTGTGGGCCAAGAGGGGGGACTGGGAGCCCAAGCCCCCTTCCAGCCCAAGGACCCAGTT 594
QY 791 TTCTGCCCTTGTCACTCCTGCCCTGTCTGTCTGCTGATGCTCCTTGTCCAAATTTGGAGCACAGCCTG 850
Db 595 TTCTGGCCTTGTCTCTTCCCTGCCCTGGTGTCCCTGCCAGAAATGGAGCACGGCCTG- 654
QY 851 GGGCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCAGTGTCCAACCAAGAACCG 910
Db 655 GGGACCTGCTCGACCACTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAGAACCG 714
QY 911 ATTCTGCCAACTGGAGATCCAAACGCGCCCTGTGTCTGCCAGACCCCTGCCCTGGCAGCCAG 970
Db 715 CTCTGCGGACTGGAGACCCAGCGCCGCTGTGCCCTGTCCAGGCCCTGCCACCCCTCCAG 774
QY 971 GAGCCACAGCTCATGGAACAGTGTCTTCTA-AGGCCAACTGGGATGCGGATACAGGGGCC 1029
Db 775 GGTGCGAGTCCACAAACAGTGTCTTAGAGCCGGGCTGGGAATGGGACACGGTGTG 834
QY 1030 TGCCATCCTCAGCAATGACCCCTAGGACCAAGCCCTGGACTGCTGGTAGATGCT-CTTCT 1088
Db 835 CACCATCCCCAGCTGGTGGCCCTGTGCCCTGGGCCCTGGGCTGGGATGGAGATGGTCCGTC 894
QY 1089 CCATGCTCTTGGTGCAGTTAACTGTCTGCTGCTTGGATTCACTGTGTAGAGC 1139
Db 895 CCAGGCCCTTGGCTGCAGGCAACACTTTAGCTTGGGTCCACCATGCAGAAC 945

RESULT 8

US-10-137-866-319
; Sequence 319, Application US/10137866
; Publication No. US20030129689A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C151
; CURRENT APPLICATION NUMBER: US/10/137,866
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974


```
RESULT 11
US-10-146-788-319
; Sequence 319, Application US/10146788
; Publication No. US20030129693A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C322
; CURRENT APPLICATION NUMBER: US/10/146,788
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-788-319

Query Match 32.9%; Score 561.4; DB 14; Length 1266;
Best Local Similarity 78.0%; Pred. No. 4e-165;
Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

QY 243 GGTGACATGAGGGGCGACCCACTGATCCATCTTCTGGCCACTTCTTCCCTTCTTCTC 302
Db 4 GGGACATGAGAGGCGACACCCGAGACCCACCTCTTCCCTTCTTCTTCTTCTC 63

QY 303 TCAATGGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCCACCCAG 362
Db 64 TCAAAGGTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCCCTGGCCACCTCCCCGA 123

QY 363 TGCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGCTGTAAGTGTGTGCACGG 422
Db 124 TGCCCGTGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGC 183

QY 423 AGSGTGGGGAGTCTGCGACCACTGATGCTGCGACCCCGAGCCAGGGCCCTGGTTGT 482
Db 184 CGSGTGGGGAGCCCTGCGACCACTCCACGCTGCGACGCGAGCCAGGGCCCTGGTCTGC 243

QY 483 CAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTTCTTGGATGAGGATGACGGT 542
Db 244 CAGCCCGGGCAGGACCCGCTGGCGGGGGCCCTGTGCTTGGCAGAGGAGCAGCAGC 303

QY 543 AGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGGTGAACGGCCGCTGTATCGGGAAAGGGAGACCTTCCAGCCCTACTGCAGC 363

QY 603 GTCCTGTGCCGCTGTGATGACGGTGGCTTACCTGCTGCCGCTGTGTCAGTGAAGATGTG 662
Db 364 ATCCGCTGCCGCTGCGAGACGGCGGCTTACCTGCTGGTGGCTGTGTCAGCGAGGATGTG 423

QY 663 CGGCTGCCAGCTGGGACTGCCCCACGCCCAAGAGATACAGGTGCCAGGAAAGTGTCTGC 722
Db 424 CGGCTGCCAGCTGGGACTGCCCCACGCCCAAGAGGTCGAGGTCTTGGCAAGTGTCTGC 483

QY 723 CCGAGTGGGTATGTACCAAGGAGTGACACCGGCGGATCCAGCGCTCCACGGCGCAAGGA 782
Db 784 TCAAAGGTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCCCTGGCCACTTCTTCTC 123
```


Db 784 ACGGTGTCCACCATCCCGAGCTGGTGGCCCTGTGCCTGGGCCCTGGGCTGATGGAAGATG 843
QY 1082 CT-CTTCTCCATGCTCTTTGGCTGCAGTTAACTGTCTCTGCTTGGATTCACTGTGTAGAGC 1139
Db 844 GTCCGTGCCAGGCCCTTGGCTGCAGGCAACACTTTAGCTTGGTCCACCATGCAGAAC 902

RESULT 14

US-10-028-072-319
; Sequence 319, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028.072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327

; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24

APPLICANT: Smith,Victoria
APPLICANT: Stewart,Timothy A.
APPLICANT: Tumas,Daniel
APPLICANT: Watanabe,Colin K
APPLICANT: Wood,William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C17
CURRENT APPLICATION NUMBER: US/10/121,049
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 319
LENGTH: 1266
TYPE: DNA
ORGANISM: Homo Sapien
US-10-121-049-319

Query Match 32.9%; Score 561.4; DB 15; Length 1266;
Best Local Similarity 78.0%; Pred. No. 4e-165;
Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;
QY 243 GGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCCTTCCCTCTGCCTTC 302
Db 4 GGGGACATGAGAGGCACACCGAAGACCCACCTCCTGGCCCTTCTCCCTCCTCTGCCTCCTC 63
QY 303 TCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACTGTCTTGGACACCAACCCAG 362
Db 64 TCAAAGTGCGTACCCAGCTGTGCCGACACCATGTACTGCCCTGGCCACCTCCCCGA 123
QY 363 TGCCACAGGGGGTACCCCTGTGTGGATGGCTGTGGCTGTAAAGTGTGTGCACGG 422
Db 124 TGCCCGTGGGAGTACCCCTGTGTGGATGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 183
QY 423 AGGCTGGGGAGTCTCGACCACTGTCATGTCTGCGACCCAGCCAGGGCCCTGTTGT 482
Db 184 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACCCAGCCAGGGCCCTGTTCTGC 243
QY 483 CAGCCTGGGGCAGGGCCCTGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT 542
Db 244 CAGCCCGGGCAGGACCCGGTGGCCGGGGCCCTGTGCCCTTGGCAGAGGACGACAGC 303
QY 543 AGCTGTGAGGTGAATGGCCGCGAGTACCTGGATGGAGAGACCTTTAAACCCAAATTGCAGG 602
Db 304 AGCTGTGAGGTGAACGGCCGCCCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGC 363
QY 603 GTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCCTGTGCAGTGAGGATGTG 662
Db 364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTACCTGCCGTGCCGTGTGCAGCGAGGATGTG 423
QY 663 CGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGAATACAGGTGCCAGGAAAGTGTGC 722
Db 424 CGGCTGCCAGCTGGGACTGCCCCACCCACCCAGGAGGTGAGGTCTCTGGGCAAGTGTGC 483
QY 723 CCGAGTGGGGTATGTGACCAAGGAGTGAACCCGGCGATCCAGCGCTCCACGGCCGAAGGA 782
Db 484 CCTGAGTGGGTGTGCGGCCAAGGAGGGGACTGGGGACCCAGCCCCCTTCCAGCCCCAAGGA 543
QY 783 CACCAACTTTCTGCCCTTGTCACTCCTCCTCTGCTGATGCTCTTGTCCAAATTGGAGC 842
Db 544 CCCCAGTTTCTGGCCCTTGTCTCTCCCTGCCCTTGGTGTCCCTGCCAGAAATGGAGC 603
QY 843 ACAGCTGGGGCCCTGTCTAACCACTGTGGGTGGGCAATAGCCACCCGAGTGTCCAAC 902
Db 604 ACGGCTGGGGACCTGTCTGACACCTGTGGGTGGGCAATGGCCACCCGGGTGTCCAAC 663
QY 903 CAGAACCGATTCTGCCAACTGGAGATCCAACGCCGCTGTGTCTGCCAGACCCCTGCCTG 962
Db 664 CAGAACCGCTTCTGCCGACTGGAGACCCAGCGCCGCTGTGCTGTCCAGGCCCTGCCCA 723
QY 963 GCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTA-AGGCCAACTGGGGATGCGGAT 1021

Db 724 CCCTCAGGGGTGCGAGTCCACAAAACAGTGCCTTCTAGAGCGCGGCTGGGAATGGGGAC 783
QY 1022 ACAGGGCTGCCATCCTCAGCAAAATGACCCCTAGGACCCAGGCGCTGGACTGCTGGTAGATG 1081
Db 784 ACGGTGCCACCATCCCCAGCTGGTGGCCCTGTGCCTGGGCCCTGGGCTGATGGAAGATG 843
QY 1082 CT-CTTCTCCATGCTCTTGGCTGCAGTTAACTGTCTGCTTGGATTCACTGTGTAGAGC 1139
Db 844 GTCCGTGCCAGGCCCTTGGCTGCAGGCAACACTTTAGCTTGGTCCACCATGCAGAAC 902

Search completed: July 28, 2003, 21:31:52
Job time : 425.038 secs



GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 18:02:17 ; Search time 3028.68 Seconds
(without alignments)
9133.322 Million cell updates/sec

Title: US-10-010-408-1
Perfect score: 1708
Sequence: 1 GAGCGTTCGATCTCCAGAG.....GCCTAGATAAACACCCAAA 1708

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				EST:*			
1:	em_estba:*	10:	gb_est2:*	17:	gb_gss:*	24:	em_gss_mus:*
2:	em_esthum:*	11:	gb_htc:*	18:	em_gss_hum:*	25:	em_gss_other:*
3:	em_estin:*	12:	gb_est3:*	19:	em_gss_inv:*	26:	em_gss_pro:*
4:	em_estmu:*	13:	gb_est4:*	20:	em_gss_pln:*	27:	em_gss_rod:*
5:	em_estov:*	14:	gb_est5:*	21:	em_gss_vrt:*		
6:	em_estpl:*	15:	em_estfun:*	22:	em_gss_fun:*		
7:	em_estro:*	16:	em_estom:*	23:	em_gss_mam:*		
8:	em_htc:*			24:	em_gss_mus:*		
9:	gb_est1:*			25:	em_gss_other:*		
10:	gb_est2:*			26:	em_gss_pro:*		
11:	gb_htc:*			27:	em_gss_rod:*		
12:	gb_est3:*						
13:	gb_est4:*						
14:	gb_est5:*						
15:	em_estfun:*						
16:	em_estom:*						
17:	gb_gss:*						
18:	em_gss_hum:*						
19:	em_gss_inv:*						
20:	em_gss_pln:*						
21:	em_gss_vrt:*						
22:	em_gss_fun:*						
23:	em_gss_mam:*						
24:	em_gss_mus:*						
25:	em_gss_other:*						
26:	em_gss_pro:*						
27:	em_gss_rod:*						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length	DB ID	Description		
c 1	725.4	42.5	762	14	BQ195526	BQ195526 UI-R-CN1-	
2	511.2	29.9	1058	14	BM805088	BM805088 AGENCOURT	
3	501.8	29.4	979	14	BQ279131	BQ279131 AGENCOURT	
4	494.6	29.0	886	13	BI822142	BI822142 603039845	
5	440.6	25.8	916	13	BI457367	BI457367 603185689	
6	436.4	25.6	750	13	BM043988	BM043988 603620978	

7	432.4	25.3	792	12	BF138093	BF138093 601785003
8	430.2	25.2	651	13	BI457141	BI457141 603185392
9	429.4	25.1	966	12	BF141695	BF141695 601790752
10	425.8	24.9	940	14	BQ937887	BQ937887 AGENCOURT
11	421.2	24.7	1006	14	BM921531	BM921531 AGENCOURT
12	417.8	24.5	1073	14	BQ073722	BQ073722 AGENCOURT
13	411.6	24.1	790	13	BM046275	BM046275 603626068
14	410.2	24.0	800	13	BI826781	BI826781 603077268
15	408.2	23.9	928	13	BI161474	BI161474 602864871
16	399.2	23.4	888	13	BI825652	BI825652 603072631
17	385.2	22.6	537	14	BQ560868	BQ560868 H4067A01-
18	375.2	22.0	1022	14	BQ952960	BQ952960 AGENCOURT
19	366	21.4	620	13	BG928868	BG928868 HNC57-1-D
20	361.2	21.1	1166	13	BM543799	BM543799 AGENCOURT
21	349.6	20.5	749	9	AL555144	AL555144 AL555144
22	346.8	20.3	488	9	AA754979	AA754979 vu55g08.r
23	336.4	19.7	1251	14	BQ961357	BQ961357 AGENCOURT
24	335.8	19.7	424	10	BB849097	BB849097 BB849097
25	321.6	18.8	489	14	BM751866	BM751866 K-EST0028
26	318.4	18.6	426	9	AA717584	AA717584 vt97c08.r
27	304.6	17.8	452	10	BE481184	BE481184 166529 BA
28	297.8	17.4	933	14	BQ278961	BQ278961 AGENCOURT
29	234.4	13.7	657	13	BM488499	BM488499 pgm2n.pk0
30	231.4	13.5	418	9	AI225477	AI225477 ue88b01.y
31	226.4	13.3	618	12	BG538695	BG538695 602566932
32	223.6	13.1	792	13	BI823598	BI823598 603040962
33	223.2	13.1	405	12	BG900069	BG900069 HOA51-1-A
34	209.6	12.3	380	12	BG900020	BG900020 HOA48-1-G
35	209.4	12.3	823	14	BQ067127	BQ067127 AGENCOURT
36	209.2	12.2	668	9	AL555143	AL555143 AL555143
37	199.4	11.7	401	12	BF849398	BF849398 CM3-EN007
38	180	10.5	369	10	BB792544	BB792544 BB792544
39	179.2	10.5	345	10	BB222795	BB222795 BB222795
40	178.4	10.4	337	10	BB319151	BB319151 BB319151
41	169.4	9.9	307	10	BB220676	BB220676 BB220676
42	167.2	9.8	313	10	BB367824	BB367824 BB367824
43	164	9.6	315	10	BB374499	BB374499 BB374499
44	161.8	9.5	697	12	BF056642	BF056642 7k20f09.x
45	154.4	9.0	948	9	AL542806	AL542806 AL542806

ALIGNMENTS

RESULT 1
BQ195526/c
LOCUS BQ195526 762 bp mRNA linear EST 30-APR-2002
DEFINITION UI-R-CN1-cmq-k-07-0-UI.s1 UI-R-CN1 Rattus norvegicus CDNA clone
UI-R-CN1-cmq-k-07-0-UI 3', mRNA sequence.
ACCESSION BQ195526
VERSION BQ195526.1 GI:20371077
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 762)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dr track served to identify it as a clone from the normalized cervix library cDNA Library Preparation: M.B. Soares Lab
Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

source

Location/Qualifiers

```
1. .762
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CN1-cmq-k-07-0-UI"
/clone_lib="UI-R-CN1"
/dev_stage="adult"
```

/lab_host="DH10B (Life Technologies)"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CN1
 library is a subtracted library derived from the following
 pool of seven normalized rat libraries: normalized rat
 seminal vesicles, normalized rat penis, normalized rat
 bladder, normalized rat cervix, normalized rat brown
 adipose, normalized rat fundus, and normalized rat
 salivary gland. It was constructed according to the
 procedure described by Bonaldo, Lennon & Soares (Genome
 Research Genome 6: 791-806, 1996). For construction of
 the CN1 library, plasmid DNA from the pool of normalized
 libraries was electroporated into competent bacteria for
 the production of single-stranded circular DNA. This was
 then used as a tracer in a subtractive hybridization with
 a driver (PCR amplified inserts from a plasmid DNA template
 preparation) comprising: a) a pool of about 34,000 clones
 from the Rat Unigene Set corresponding to plates R-5-AA-NN
 excluding plates R-5-MM and MN. This pool represented 40%
 of the final driver population. b) a pool of about 29,000
 clones from subtracted libraries CA0 and CAL corresponding
 to plates R-CA0-AWV through R-CA0-AXS, R-CA0-AZX through
 R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS,
 R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through
 R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA,
 R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through
 R-CA0-BOD, R-CA0-BPA through R-CA0-BPG, R-CA0-BBA through
 R-CA0-BDJ, R-CA0-BHZ through R-CA0-BJF, R-CA0-BJR,
 R-CA0-BJT through R-CA0-BKB, R-CA0-BKD, R-CA0-BKF,
 R-CA0-BKI, R-CA0-BKT, R-CA0-BLF, R-CA0-BLH through
 R-CA0-BLN, R-CA0-BLS, R-CA0-BLU-V, R-CA0-BNR, and
 R-CA0-BLE. The resulting pool represented 20% of the
 final driver population. c) a pool of about 15,000 clones
 from non-normalized libraries CS0s, CT0s, CU0s, CW0s, CX0s
 and normalized libraries CS0, CT0, CU0, CW0, and CX0
 corresponding to plates R-CS0s-CBD through R-CS0s-CBO,
 R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through
 R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN
 through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV
 , R-CS0-BVM, R-CT0-BTW through R-CT0-BUP, R-CT0-BVN,
 R-CU0-BUQ through R-CU0-BVL, R-CW0-BVV through R-CW0-BWP,
 R-CX0-BXN through R-CW0-BXO, R-CX0-BWQ through R-CX0-BXM.
 The resulting pool represented 5% of the final driver
 population. d) a pool of about 5,000 clones (1,000 from
 non-normalized eye library CV0 and 4,000 from normalized
 eye library CV1) corresponding to plates R-CV0-BRH through
 R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through
 R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool
 represented about 5% of the final driver population. e) A
 pool of about 10,000 clones from subtracted library BS2,
 BV0 and BV0p (7-9.5 kb cDNA library fraction from rat
 whole embryo), and BX0 (0.5-7kb cDNA library fraction from
 rat whole embryo) corresponding to plates R-BS2-BDB
 through R-BS2-BFB, R-BV0-ANK through R-BV0-ANR, R-BV0p-AOI
 through R-BV0p-AOX, and R-BX0-AQY through R-BX0-ASH. The
 resulting pool represented 5% of the final driver
 population. f) a pool of about 7,000 clones from the
 seven non-normalized libraries that make up the tracer

including CY0, CZ0, DA0, DB0, DC0, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYI, R-CZ0-BZB-C, R-DA0-BYJ through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ through R-DB0-BZA, R-DC0-BZI through R-DC0-BZQ, R-DD0-CAY through R-DD0-CBA, R-DD0-BZR through R-DD0-CAA, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CN0, that makes up the tracer. The corresponding plates are R-CN0-BKW through R-CN0-BLD, R-CN0-BLG, R-CN0-BLP through R-CN0-BLR, R-CN0-BLT, R-CN0-BLW-X, R-CN0-BMB, and R-CN0-BMF through R-CN0-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CN0 pool corresponding to the following addresses: bkx-a-09-0-UI, bkx-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI, bkx-d-01-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI, bkx-h-12-0-UI, bkx-a-05-0-UI, bkz-a-06-0-UI, bkz-a-11-0-UI, bkz-c-06-0-UI, bkz-c-09-0-UI, bkz-g-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-e-05-0-UI, bld-1-08-0-UI, bld-f-02-0-UI, blq-h-04-0-UI, blr-a-05-0-UI, blt-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CN0 clone (corresponding to the address bkz-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG_LIB=UI-R-CN1
TAG_TISSUE=cervix
TAG_SEQ=GACCA"

BASE COUNT	177 a	176 c	198 g	211 t
ORIGIN				

```
Query Match 42.5%; Score 725.4; DB 14; Length 762;
Best Local Similarity 99.2%; pred. NO. 3.9e-186;
Matches 729; Conservative 0; Mismatches 6; Indels 0;
```

QY	974	CCACAGCTCATGGAA	CAGTCTTTCTAAGGCCAACTGGGGATCGGGATACAGGGCCTGCC	1033
Db	762	CCACAGCTCATGGAA	CAGTCTTTCTAAGGCCAACTGGGGATCGGGATACAGGGACTGTC	703
QY	1034	ATCCTCAGCAAAATGA	CCCTAGGACCAGGCCCTGGACTGCTGGTAGATGCTCTTCTCCATG	1093
Db	702	ATCCTCAGCAAAATGA	CCCTAGGACCAGGCCCTGGACTGCTGGTAGATGCTCTTCTCCATG	643
QY	1094	CTCTTGGCTGCAGTT	AACTGTCTCTTGGATTCACTGTGTAGAGCCACTGAGCGATCCC	1153
Db	642	CTCTTGGCTGCAGTT	AACTGTCTCTTGGATTCACTGTGTAGAGCCACTGAGCGATCCC	583
QY	1154	TGCTCTGTCTGAGGT	AGGGCGGAGCAGGTGACCAAGCTCCAGTTCTCTGGTTCAGCCTGGAA	1213
Db	582	TGCTCTGTCTGAGGT	AGGGCGGAGCAGGTGACCAAGCTCCAGTTCTCTGGTTCAGCCTGGGA	523
QY	1214	TTCTGGGTTCTCCTGG	CTCATTCCTCAAAACATCCCTGTACAAAAAGGACAACCAAAAAG	1273
Db	522	TTCTGGGTTCTCCTGG	CTCATTCCTCAAAACATCCCTGTACAAAAAGGACAACCAAAAAG	463
QY	1274	ACCTTTAAACCTTAGG	CTATACTGGGCAAAACCTGGCCACCCTGCTGGGATTAAGGTCAATG	1333
Db	462	ACCTTTAAACCTTAGG	CTATACTGGGCAAAACCTGGCCACCCTGCTGGGATTAAGGTCAATG	403
QY	1334	TTAGGACCAGACAGCA	GATTGCCTGAAACTTCCAATTCCCTTCTTGGACTTCTGTATGCT	1393
Db	402	TTAGGACCAGACAGCA	GATTGCCTGAAACTTCCAATTCCCTTCTTGGACTTCTGTATGCT	343
QY	1394	TGTCCCCAAAAGATGA	TGAATGAATCGTAAGTGTACCTTCCCTGACCTGAGAACACCCCTG	1453
Db	342	TGTCCCCACAGATGA	TGAATGAATCGTAAGTGTACCTTCCCTGACCTGAGAACACCCCTG	283
QY	1454	CCTGCTCGGGAAGTAT	TTCAGGGGCGAGAAATTCTCTGTGAACATGAAGAGATGAATCACACT	1513

Db 282 CCTGCTCGGAAGTATTTCAGGGGAGAAATCTCTGTGAACATGAAGAGATGAATCACACT 223
QY 1514 GTCTTAAGAAATTCCTGAAAGTCACAGGAACCTTGAGCTTTGATTTTCAGGAATGCACAT 1573
Db 222 GTCTTAAGAAATTCCTCAAAGTCCAGGAACCTTGAGCTTTGATTTTCAGGAATGCACAT 163
QY 1574 CTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTGGCAGGCCCTTTCTCT 1633
Db 162 CTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTGGCAGGCCCTTTCTCT 103
QY 1634 TCAGCATGAGAAAGACAAGGGACACAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCTA 1693
Db 102 TCAGCATGAGAAAGACAAGGGACACAGAGTACTCTCCTCTGGAGGACTAGTCCAGCCTA 43
QY 1694 GAATAAACACCCAAA 1708
Db 42 GAATAAACACCCAAA 28

RESULT 2
BM805088 1058 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6490429 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5725937
DEFINITION 5', mRNA sequence.
ACCESSION BM805088
VERSION BM805088.1 GI:19121911
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1058)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM12717 row: j column: 18
High quality sequence stop: 697.
Location/Qualifiers
1. .1058
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5725937"
/clone_lib="NIH_MGC_125"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT 139 a 423 c 294 g 200 t 2 others
ORIGIN
Query Match 29.9%; Score 511.2; DB 14; Length 1058;
Best Local Similarity 79.3%; Pred. No. 8.1e-128;
Matches 606; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
QY 242 CGGTGACATGAGGGGCAGCCCCACTGATCCATCTTCTGGCCACTTCCTTCCTGCTCTCT 301
Db 2 CAGGGGAATGAGAGGCACACCCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCTCTCT 61

QY 302 CTCAATGGTGTGTGCCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTTGGACACCCCA 361
Db 62 CTCNAAGGTGCGTACCCAGCTGTCCCCGACACCATGTACCTGCCCTTGGCCACCTCCCG 121
QY 362 GTGCCACAGGGGGTACCCCTGCTGGATGGCTGTGGCTGTCTGTAAAGTGTGTGCACG 421
Db 122 ATGCCCGCTGGAGTACCCCTGCTGGATGGCTGTGGCTGTCTGCTGCCGGTATGTGCACG 181
QY 422 GAGCTGGGGGAGTCTCTCGGACCACTGCATGTCTGGGACCCCAAGCCAGGCGCTGGTTG 481
Db 182 GCGGTGGGGAGCCCTGCGCACTCCACGTCTGCGACGCCAGCCAGGCGCTGGTCTG 241
QY 482 TCAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGG 541
Db 242 CCAGCCCGGGCAGGACCCGCTGGACGGGGGCCCTGTGCCTCTTGGCAGAGGACGACG 301
QY 542 TAGCTGTGAGGTGAATGGCCGACAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG 601
Db 302 CAGCTGTGAGGTGAACGGCCGCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAG 361
QY 602 GGTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGCAGTGAGGATGT 661
Db 362 CATCCGCTGCCGCTGCGAGGACGCGGCTTACCTGCGTGGCTGTGCAGCGAGGATGT 421
QY 662 GCGCTGCCCCAGCTGGGACTGCCCCACGCCCAAGAGATAACAGGTGCCAGGAAAGTGTG 721
Db 422 GCGGTGCCCCAGCTGGGACTGCCCCACGCCCAAGAGGAGGTGCGAGGTCTGGCAAGTGTG 481
QY 722 CCCGAGTGGGTATGTGACACGAGGAGTGACACCGGGATCCAGCGCTCCACGGCGCAAGG 781
Db 482 CCCTGAGTGGGTGTGCGGCCAAGAGGGGGACTGGGGACCCAGCCCCCTTCCAGCCCCAAGG 541
QY 782 ACACCAACTTTCTGCCCTTGTCACTCCTGCTGTGCTGTGATGCTCTTGTCCAAATGGAG 841
Db 542 ACCCAGTTTCTGGCCTTGTCTCTTCTCCCTGCCCCCTGGTGTCCCTGCCAGAATGGAG 601
QY 842 CACAGCTGGGGCCCCCTGTCAACCACTGTGGGTGGGCATAGCCACCCGAGTGTCCAA 901
Db 602 CACGGCTGGGGACCCCTGCTCGACCACTGTGGGTGGGCATGGCCACCCGGGTGTCCAA 661
QY 902 CCAGAACCGATTCTGCCAACTGSAGATCCAAACGCCCTGTGTCTGTCCCAGACCCCTGCC 961
Db 662 CCAGAACCGCTTCTGCCGACTGSAGACCCAGCGCCCTGTGTCTGTCCCAGGCCCTGCC 721
QY 962 GGCAGCCAGGAGCCACAGCTCATCGAACAGTGTCTTCTAAGGCC 1005
Db 722 ACCTCCAGGGGTGCGAGTCCACAAAACAGTGCCTTCTAGAGCC 765

RESULT 3
BQ279131 979 bp mRNA linear EST 07-MAY-2002
LOCUS AGENCOURT_7046721 NIH_MGC_107 Homo sapiens cDNA clone IMAGE:5805819
DEFINITION 5', mRNA sequence.
ACCESSION BQ279131
VERSION BQ279131.1 GI:20489339
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 979)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

QY	495	GGCCCTGGCGCCCATGGGGCTGTGTGCTCTCTTGATGAGGATGACGGTAGCTGTAGGTG	554
Db	241	GACCCCTGTGGACGGGGCCCTGTGCTCTTGGCAGAGGACAGACAGCTGTGAGGTG	300
QY	555	AATGGCCGACGGTACCTGGATGGAGAGACACCTTTAAACCCCAATTGCAGGGTCTGTGCCGC	614
Db	301	AACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGC	360
QY	615	TGTGATGACGGTGGCTTACCTGCCCTGCCGCTGTGCAGTGAGGATGTGCGGCTGCCCAGC	674
Db	361	TGCGAGGACGGGGCTTCACTGCTGCGGCTGTGCAGCGAGGATGTGCGGCTGCCCAGC	420
QY	675	TGGGATGCCCCACGCCCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCCGAGTGGGTA	734
Db	421	TGGGACTGCCCCCACCCACGAGGAGGTCGAGGTCTCTGGCAAGTGTGCCCTGAGTGGGTG	480
QY	735	TGTGACCAAGGAGTGACACCGCGGATCCAGCGCTCCACGGCGGAAGGACACCAACTTCT	794
Db	481	TGCG-GCAAGAGGGGACTGGGGACCCAGCCCCCTTCCAGCCCCAAGGACCCAGTTTCT	539
QY	795	GCCCTTGACACTCCTGCTCTGCTGATGCTCCTTGTCCTCAATTTGGAGCACAGCCTGGGC	854
Db	540	GCCCTTGCTCTCTCCCTGCCCCCTGGTGTCCCCTGCCAGATGGAGCACGGCCTGGGA	599
QY	855	CCCTGCTCAACCACTGTGGCTGGCATAGCCACCCGAGTGTCCAACCAAGAACCGATTTC	914
Db	600	CCCTGCTGACCACTGTGGCTGGCATGGCCACCCGCGGTGTCCAACCAAGAACCGCTTC	659
QY	915	TGCCAAGTGAGATCCAAAGCCGCTGTGTCTGCCCCAGACCTGCTGGCAGCCAGGAGC	974
Db	660	TGCCGAAGTGAGACACAGCCGCGCTGTGCTGTCCAGGCCCTGCCACCCCTCCAGGGT	719
QY	975	CACAGCTCATGGAACAGTGTCTTCTA-AGGCCAACTGGGGATGGGGATACAGGCGCTGCC	1033
Db	720	CGCAGTCCACAAAACAGTGCCCTTAGAGCCGGGCTGGGAATGGCGACACGGTGTCAACC	779
QY	1034	ATCCTCAGCAAAATGACCCTAGGACAGGCGCTGGACTG-CTGGTAGATGCTCTTCTCCAT	1092
Db	780	ATCCCCAGCTGGTGGCCTGTGCTGTGCTGCTGGCTGGCTGACTGGAAGATGTCGCTGCCA	839
QY	1093	GCTCTGGCTGCAGTTAACTGTCTGCTTGGATTCACTGTG	1133
Db	840	GGCCTTGGGTGCAGGCAACACTTAGCTAGGGTTCACCATG	880
RESULT 5	BI457367		
LOCUS	BI457367	916 bp	mRNA linear EST 21-AUG-2001
DEFINITION	603185689F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5258398 5', mRNA sequence.		
ACCESSION	BI457367		
VERSION	BI457367.1	GI:15248023	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 916)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Ling Hong/Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	plate: L1CM1870 row: a column: 23		
	High quality sequence stop: 778.		

FEATURES		Location/Qualifiers	
source		1..916	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5258398"	
		/clone_lib="NIH_MGC_42"	
		/tissue_type="epithelioid carcinoma cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. "	
BASE COUNT		136 a	305 c 304 g 170 t 1 others
ORIGIN			
Query Match		25.8%; Score 440.6; DB 13; Length 916;	
Best Local Similarity		77.7%; Pred. No. 1.2e-108;	
Matches 594;		Conservative 0; Mismatches 165; Indels 5; Gaps 5;	
QY	279	GCCACTTCTCTCTGCTCTCTCAATGGTGTGCCCCAGCTGTGCCGGACACCCCTGT	338
Db	.1	GCCCTCTCCCTCTGCTCTCTCAAAGTGCCTACCCAGCTGTGCCCGACACCATGT	60
QY	339	ACCTGTCTTGGACACACCCACGTGCCACAGGGGTACCCCTGGTGGTGGCTGTGT	398
Db	61	ACCTGCCCTGG-CACCTCCCCGATGCCCGCTGGGAGTACCCCTGGTGGTGGCTGT	119
QY	399	GGCTGCTGTAAAGTGTGTGCACGGAGGCTGGGGGAGTCTCGGACCACTGATGTCTGC	458
Db	120	GGCTGCTGCCGGTATGTGCACGGCGGCTGGGGGAGCCCTGCGACCAACTCCACGTCTGC	179
QY	459	GACCCAGCCAGGGCCTGGTTGTACGCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTG	518
Db	180	GACGCCAGCCA-GGCCTGGTCTGCCAGCCCGGGCAGGACCCCGTGGACGGGGGCCCTG	238
QY	519	TGCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGA	578
Db	239	TGCCTCTTGGCAGAGGACGACAGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGG	298
QY	579	GAGACCTTTAAACCCAATTGCAGGGTCTGTGCCGCTGTGATGACGGTGGCTTCACTGC	638
Db	299	GAGACTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGGCTTCACTGC	358
QY	639	CTGCCGCTGTGCAGTGAGGATGTGCGGCTGCCAGCTGGGACTGCCACGCCCAAGAGA	698
Db	359	GTGCCGCTGTGCAGCGAGGATGTGCGGCTGCCAGCTGGGACTGNCCTTACCCAGGAGG	418
QY	699	ATACAGGTGCCAGGAAAGTGTGCCCGAGTGGGTATGTGACCAAGGAGTGACACCCGGG	758
Db	419	GTCGAGGTCTCTGGCAAGTGTGCCCTGAGTGGGTGTGCGGCCAAGGAGGGGACTGGGG	478
QY	759	ATCCAGGCTCCACGGGCAAGGACACCAACTTCTGCCCTTGTCACTCTCTGCTGTCT	818
Db	479	ACCCAGGCCCTTCCAGCCCAAGGACCCCAAGTTTCTGGCCTTGTCTCTTCCCTGCCCT	538
QY	819	G-ATGCTCTTGTCCAAATTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGTGGGCT	877
Db	539	GTGTGCTCCCTGCCAGAATGGAGCACGGCCTGGGACCCCTGCTCGA-CACCTGTGGGCT	597
QY	878	GGGCATAGCCACCCGAGTGTCCAACCAAGAACCGATTCTGCCAACTGGAGATCCAACGCCG	937
Db	598	GGGCATGGTCAACCCGGGTGTCCAACAGAACCGGTTCTGCCGACTGGAGACCCAGGCCG	657
QY	938	CC-TGTGTCTGCCCAGACCCCTGCCCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGT	996
Db	658	CCTTGTGCTGTCTCAGGCCCTGACCAAGCCTCCAGGGGTCCAGAGTCCACACACAGTGT	717
QY	997	TCTAAGGCCCAACTGGGGATGGGATACAGGGCCTGCCATCTCTCA	1040

Db	718	CTAGAGCCGGCTGGGACTGGGACACACGGTGTCCACCATTC	761
RESULT 6			
BM043988			
LOCUS			
DEFINITION	603620978F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5446794 5', mRNA sequence.	750 bp	linear EST 07-NOV-2001
ACCESSION	BM043988		
VERSION	BM043988.1	GI:16773255	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 750)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM1929 row: k column: 19 High quality sequence stop: 714.		
FEATURES			
source	1. .750		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5446794"		
	/clone_lib="NIH_MGC_40"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	112 a 267 c 246 g 125 t		
ORIGIN			
	Query Match 25.6%; Score 436.4; DB 13; Length 750;		
	Best Local Similarity 79.0%; Pred. No. 1.6e-107;		
	Matches 569; Conservative 0; Mismatches 146; Indels 5; Gaps 4;		
QY	243	GGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTGCTC	302
Db	10	GGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTGCTC	69
QY	303	TCAATGGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCC	362
Db	70	TCAAAGGTGCTACCCAGCTGTGCCCGACACCATGTACCTGCCCTTGGCCACCT	129
QY	363	TGCCCCACAGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGCTGTAAGTGTG	422
Db	130	TGCCCCGTGGGAGTACCCCTGGTGTGCTGGATGGCTGTGGCTGCTGCGGGTATG	189
QY	423	AGGCTGGGGAGTCTCGACACCACTGTCGACCCCA-GCCAGGGCCTGGTTG	481
Db	190	CGGCTGGGGAGCCCTGCGACCAACTCCAGCTCTGCGACGCCATGCCAGGCCCTG	249
QY	482	TCAGCCTGGGGCAGGCCCTGGCGGCCCATGGGGCTGTGTCTCTTGGATGAGGATG	541
Db	250	CCAGCCCCGGGACAGGACCCGGTGGACGGGGGGCCCTGTGCCTCTTGGCAGAGG	309

QY	542	TAGCTGTGAGGTGAATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAAT	601
Db	310	CAGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCAC	369
QY	602	GGTCTGTGCTGCTGTGATGACGGTGGCTTACCTGCCTGCCCTGTGCAGTGAGG	661
Db	370	CATCCGCTGCCGCTGGAGGACGGCGGCTTACCTGCGTGCCTGTGCAGCGGAG	429
QY	662	GCGGCTGCCAGCTGGGACTGCCACGCCGCCAAGAGATACAGTGCCAGGAAGTG	721
Db	430	GCGGCTGCCAGCTGGGACTGCCACGCCGCCAAGAGGTCGAGGTCTGGCAAGTG	489
QY	722	CCCCGAGTGGGTATGTACACGAGGAGTACACCGCGGATCC-AGCGCTCCACG	780
Db	490	CCCTGAGTGGGTGTGCGGCCAAGGAGGGGACTGGGGACCCAAAGCCCTTCC	549
QY	781	GACACCACTTGTGCCCCTGTGACTCTCTGCTGCTGCTGATGCTCTGTCCTG	840
Db	550	GACCCCACTTGTGCCCCTGTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTG	609
QY	841	GCACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCACCCGAG	900
Db	610	GCACGCGCTGGGACCTGTCTGA-CACCTGTGGCTGGGCATGCGCACCCGGG	668
QY	901	ACCAGAACCGATTGTGCCAACTGGAGATCCAAACGCCGCTGTGTGCTGCTG	960
Db	669	ACCAGAACCGTTCTAGCCGACTGGAGA--CCAGCGCGCTGTGCTGCTGCTG	726
RESULT 7			
BF138093			
LOCUS	601785003F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4012801 5', mRNA sequence.	792 bp	mRNA linear EST 24-OCT-2000
DEFINITION	mRNA sequence.		
ACCESSION	BF138093		
VERSION	BF138093.1	GI:10977133	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 792)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM9254 row: f column: 02 High quality sequence stop: 705.		
FEATURES			
source	1. .792		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:4012801"		
	/clone_lib="NCI_CGAP_Lu30"		
	/tissue_type="tumor, metastatic to mammary"		
	/lab_host="DH10B"		
	/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"		
BASE COUNT	193 a 204 c 205 g 190 t		
ORIGIN			

Query Match 25.3%; Score 432.4; DB 12; Length 792;
Best Local Similarity 82.8%; Pred. No. 2e-106;
Matches 596; Conservative 0; Mismatches 101; Indels 23; Gaps 8;

QY 939 CTGTGTCGCCCCAGACCCTGCCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGCTTC 998
Db 1 CTGTGTCGCCCCAGACCCTGCCTGGCAGCCAGGAGCCAGCTCATGGAACAGTGCTTC 60

QY 999 TAAGGCCA-ACTGGGGATCGGATACAGGGCCTGCCATCCTCAGCAAAATGACCCTAGGAC 1057
Db 61 TAGAGCCATTGTGGGATGTGAATACAGGGCCTGCCATTCTCAGCAAAATGTCCCTAGGAC 120

QY 1058 CAGGCCTGGACTGCTGGTAGATGCTCTTCTCCATGCTCTGGCTGCAGTTAACTGTCT 1117
Db 121 CAGGCCTGGACTGATGGTAGATGCCCTCTCCATGCTCTGGCTGCAGTTAACTGTCT 180

QY 1118 GCTTGATTCACTGTGTAGAGCCACTGAGCGATCCCTGCTGCTGTGAGTAGGGGAGC 1177
Db 181 GGGTGGATTCAGTGCCAGAGCCTCTGAGCGATCCCTGCTGCTGTGAGTGGGGAGC 240

QY 1178 AGGTGACCAGCTCCAGTTCTCTGG--TTCAGCCTGGAATCTGGGTTCTCCTGGCTCAT 1235
Db 241 AGGTGACCAGCTACATTTCTCTGATCTGACCCAGGCTTCTGGGTTCTCCTGGCTAGT 300

QY 1236 CCTCAAAACATCCCTGTACAAAAGGACAAACCAAAAGACCTTTAAACCTAGGCTATACT 1295
Db 301 CCTCAAAACTTCCCTGTATGAAAAGGACAAACCAAAAGGACCTTTAAAGCTAAGCTGTA 360

QY 1296 GGGCAAAACCTGGCCACCGTGTGGGGATAAGTCAATGTTAGG-ACCAGACAGCATG 1354
Db 361 GGGCAAGCCTGG-CACCATGCTGGGATAGTGACAGTAATAGGTACCGACGACAGATTG 419

QY 1355 CCTGAAACTTCCAATTCCTTCTTGGACTTCTGTATGCTGTGTCCTCCCAAGATGATG 1414
Db 420 CCTGAAACATCCAGTCCCTTCTTGGACTTCTATGTGCTTGTCCC-AAAGATTATGGGTG 478

QY 1415 AACTCGTAAAGTGTACCTTCCCTGACCTGAGAACACCCCTGCCTGCTCGGGAAGTATTCAGG 1474
Db 479 ACCTTGTAAAGTGTGCTTTCCTGATCTGAGAACACCCCTGCGGCT----- 523

QY 1475 GGCAGAATTCCTGTGAACATGAAGAGAT-GAATCACACTGTCCTTAAGAAATTCCTGAA 1533
Db 524 GGAAGAATTTCTGGGAACATGAAGAGATGGAATCACACTATTCTTAAGAGCGTTGGCCA 583

QY 1534 AGTCCAGGAACCTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGGAAAC 1593
Db 584 AGTCCAGGAACCTTGACCTGTGTATTTGTAACAATACACATCTTTAA-AATGCTCACAA 642

QY 1594 AGGAAGCTCCACACCTCTGGCAGGCCAGGGCCTTTCTTTCAGCATGAGAAAGACAGG 1653
Db 643 GCAAGAGTCCACACTTCTGGCAGGCCAGGGCCTTTCTTTCAGCATGAGAGACAGG 702

RESULT 8
BI457141

LOCUS BI457141 651 bp mRNA linear EST 21-AUG-2001
DEFINITION 603185392F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5258159 5', mRNA sequence.

ACCESSION BI457141
VERSION BI457141.1 GI:15247797
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 651)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1869 row: g column: 24
High quality sequence stop: 651.
Location/Qualifiers
1. .651
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5258159"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using 2AP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. |"
BASE COUNT 91 a 230 c 219 g 111 t
ORIGIN
Query Match 25.2%; Score 430.2; DB 13; Length 651;
Best Local Similarity 80.1%; Pred. No. 7.3e-106;
Matches 518; Conservative 0; Mismatches 128; Indels 1; Gaps 1;

QY 285 TCCTTCTCTGCCTTCTCTCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGT 344
Db 6 TCCTTCTCTGCCTTCTCTCAAGGTGCGTACCCAGCTGTGCCGACACCATGTACCTGC 65

QY 345 CCTTGGACACACCACCAGTGCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGC 404
Db 66 CCTTGGCCACCTCCCCGATGCCGCTGGAGTACCCCTGGTGTGGATGGCTGTGGCTGC 125

QY 405 TGTAAAGTGTGTGCACGGAGGTGGGGAGTCTCTGCGACACCTGCATGTCTGCGACCCC 464
Db 126 TGCCGGGTATGTGCACGGCGCTGGGGAGCCCTGCGACAACTCCACGTCTGCGACGCC 185

QY 465 AGCAGGGCCTGGTTTGTACAGCTGGGCGAGCCCTGGGCGCATGGGCTGTGTCTC 524
Db 186 AGCCA-GGCCTGGTCTGCCAGCCGGGCGAGGACCCGCTGACGGGGGCGCCCTGTGCCTC 244

QY 525 TTGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGACAGTACCTGGATGGAGAGACC 584
Db 245 TTGCGAGAGGACGACAGCAGCTGTGAGGTGAACGGCCGCTGTATCGGAAGGGAGACC 304

QY 585 TTTAAACCCCAATTGCAGGGTCTGTGCCGCTGTGTATGATGACGGTGGCTTCACCTGCCCTGCCG 644
Db 305 TTCCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGGCTTCACCTGCCGTGCCG 364

QY 645 CTGTGAGTGAGGATGTGCGGCTGCCAGCTGGGACTGGCCAGCCGCCAAGAGAATACAG 704
Db 365 CTGTGACGCGAGGATGTGCGGCTGCCAGCTGGGACTGCCCCACCCAGAGGGGTCCGAG 424

QY 705 GTGCCAGGAAAGTGTGCCCCGAGTGGGTATGTGACCGAGGAGTGACACCGCGCATCCAG 764
Db 425 GTCTGGGCAAGTGTGCTGCCCTGAGTGGGTGTGCGGCCAAGAGGGGGGACTGGGGACCCAG 484

QY 765 CGCTCCACGGCGCAAGGACACCAACTTTCTGCCCTTGTCACTCTCTGCCTCTGTGATGCT 824
Db 485 CCCCTCCAGCCCAAGGACCCAGTTTCTGGCCTTGTCTCTCTCCCTGCCCCCTGGTGT 544

QY 825 CCTTGTCCAAATTGGAGCACAGCCTGGGGCCCTGCTCAACCACCTGTGGGCTGGGCATA 884
Db 545 CCCTGCCCGAGAATGGAGCACGGCCTGGGACCCCTGCTCGACCACTGTGGGCTGGGCATG 604

QY 885 GCCACCCGAGTGTCCAAACCAACCGGATTCTGCCAACTGGAGATCCA 931


```

Db      605 GCCACCGGGTGTCCAAGCAGAACCGGTTCTGCCGACTGGAGACCCA 651
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
RESULT 9
BF141695
LOCUS
DEFINITION 601790752F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4021308 5',
          mRNA sequence.
ACCESSION BF141695
VERSION   BF141695.1 GI:10980735
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 966)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Tissue Procurement: Gilbert Smith, Ph.D.
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM9276 row: h column: 13
           High quality sequence stop: 613.
           Location/Qualifiers
             1. .966
               /organism="Mus musculus"
               /strain="CZECH II"
               /db_xref="taxon:10090"
               /clone="IMAGE:4021308"
               /clone_lib="NCI_CGAP_Lu30"
               /tissue_type="tumor, metastatic to mammary"
               /lab_host="DH10B"
               /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
               Site_2: Sali; transgenic model WNT-1, expression driven by
               MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
               dt. Library constructed by Life Technologies.
               Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 227 a 298 c 258 g 183 t
ORIGIN

Query Match      25.1%; Score 429.4; DB 12; Length 966;
Best Local Similarity 85.6%; Pred. No. 1.4e-105;
Matches 571; Conservative 0; Mismatches 86; Indels 10; Gaps 8;

QY      808 CTGCCTCTGCTGATGCTCCTTGTCCTCAAAATGGAGCACAGCCCTGGGGCCCTGCTCAACCA 867
Db      1 CTGCATCTGCCGATGGCCCTGTCCAAACTGGAGCACAGCCTGGGGCCCTGTTCAACCA 60
QY      868 CCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAGAACCGATTCTGCCAACTGGAGA 927
Db      61 CTGTGGGTGGGCATAGCCACCCGAGTATCCAACCAGAACCGATTCTGCCAACTGGAGA 120
QY      928 TCCAACGCCGCTGTGTCTGCCCCAGACCCCTGCCTGGCAGCCAGGACCAAGCTCATGGA 987
Db      121 TCCAGCGTCGCTGTGTCTGTCCAGACCCCTGCCTGGCATCCAGGAGCCACGGCTCATGGA 180
QY      988 ACAGTGTCTTCTAAGGCCA-AC TGGGGATGGGGATACAGGGCCTGCCATCCTCAGCAAAT 1046
Db      181 ACAGTGCCTTCTAGAGCCATTGTGGGGAFTGTAATACAGGGCCTGCCATTCTCAGCAAAT 240
QY      1047 GACCCTAGGACAGGCCCTGGACTGCTGGACTGCTTCTTCCATGCTTGGCTGCAG 1106
Db      241 GTCCCTAGGACAGGCCCTGGACTGATGGACTGATGCCCTCTCCATGCTTGGCTGCAG 300
QY      1107 TTAAGTGTCTGCTTGGATTTCACCTGTCTAGAGCCACTGAGCGGATCCCTGCTCTGTCTGAG 1166
```

```

Db      301 TTAAGTGTCTGCTGGTGGATTAGTGTCCAGAGCCTCTGAGCGATCCCTGCTGTCTGAG 360
||||||| ||||||| ||||||| ||| ||||||| ||||||| ||||||| |||||||
QY      1167 GTAGGCGGAGCAGGTGACCAGCTCCAGTTCTCTGG--TTCAGCCCTGGAATTTCTGGTTCT 1224
          || || || || || || || || || || || || || || || || || || || || || ||
Db      361 GTGGGGGAAGCAGGTGACCAGCTACATTTCTCTGGATTCTGACCCAGGCTTCTGGGTTCT 420
          || || || || || || || || || || || || || || || || || || || || || ||
QY      1225 CCTGGCTCATTTCTCAAAACATCCCTGTACAAAAGGACAACCAAAAAGACCTTTAAACC 1284
          || || || || || || || || || || || || || || || || || || || || || ||
Db      421 CCTGGCTAGTTCTCAAAACTTCCCTGTATCAAAAGGACAACCAAAAGGACCTTTAAAGC 480
          || || || || || || || || || || || || || || || || || || || || || ||
QY      1285 TAGGCTATACTGGGCAACCTGGCCACCGCTGGGGATAAGGTCAATGTAGG-ACCAG 1343
          || || || || || || || || || || || || || || || || || || || || || ||
Db      481 TAAGCTGTACT-GGCAAGCCTGGCCACCATGCTGGGATAGTGACAGTAATAGGTACCAG 539
          || || || || || || || || || || || || || || || || || || || || || ||
QY      1344 ACAGCAGATTGCTGAAACTTCCAATTCCCTTCTTGGACTTCTGTATGCTGTGCCCAA 1403
          || || || || || || || || || || || || || || || || || || || || || ||
Db      540 GCAGCAGAT--GCTGAAACATCCAGGTCCCTTCTTGGACTTCTATGTGCTGT-CCCACA 596
          || || || || || || || || || || || || || || || || || || || || || ||
QY      1404 GATGATGAATGAACCTGTAAGTGTACCTTCCCTGACCTGAGAACACCCCTGCTCGGG 1463
          || || || || || || || || || || || || || || || || || || || || || ||
Db      597 GATTATGGGTGACCT-GTAAGTGTGCC-TCCCTGATCTGAGAACACCTGSCCGGGGGC 654
          || || || || || || || || || || || || || || || || || || || || || ||
QY      1464 AAGTATT 1470
          || || || |
Db      655 ACGAAAT 661

RESULT 10
BQ937887
LOCUS
DEFINITION BQ937887 940 bp mRNA linear EST 21-AUG-2002
          AGENCOURT_8951807 NCI_CGAP_Co24 Mus musculus cDNA clone
          IMAGE:6476852 5', mRNA sequence.
ACCESSION BQ937887
VERSION   BQ937887.1 GI:22353365
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 940)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: The Cepko Laboratory
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM14017 row: n column: 21
           High quality sequence stop: 543.
           Location/Qualifiers
             1. .940
               /organism="Mus musculus"
               /strain="FVB/N"
               /db_xref="taxon:10090"
               /clone="IMAGE:6476852"
               /clone_lib="NCI_CGAP_Co24"
               /lab_host="DH10B (T1 phage-resistant)"
               /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
               Site_2: Sali; Cloned unidirectionally. Primer: Oligo dt.
               Average insert size 1.6 kb. Constructed by Life
               Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 169 a 277 c 288 g 200 t
ORIGIN

Query Match      24.9%; Score 425.8; DB 14; Length 940;
Best Local Similarity 90.9%; Pred. No. 1.3e-104;
Matches 488; Conservative 0; Mismatches 39; Indels 10; Gaps 3;
```

QY 1 GACGCTTCTGATCTCCAGAGGACCOCTGGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCC 60
|||||
Db 37 GACGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCTTGGCGAGGCTGCAGCT 96
QY 61 GCTG-GGCAGTGGCTTGAATGAGAGGTCTTTATTACTGGGAAGTGGAGGCTAAGAGGCT 119
|||||
Db 97 GCTGTGGCAGTAGCTTGGGATGAGAGGTCTTTCTTGTCTGGGAAGTGGAGGCTGAGAGCT 156
QY 120 CCTGTACG---CTTGTCTTAAGTCTTAGCACCTTGTGGTGGCTTGGGCTTCACACACTGT 176
|||||
Db 157 CCTGTACAGGCTCCTGTCTTAAGTCTTGGACCTTGGGCTGGGCTTCACACACTGT 216
QY 177 CAGACACCTTCGTGGTGGCTTCCAGGCGCTCACCTTCAAGTTTGAAGCTGGCTCCACAAG 236
|||||
Db 217 CAGACACCTTCTTGGTGGCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAG 270
QY 237 GGACAGGCTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCCCTTCCCTGTC 296
|||||
Db 271 GGACAGGCTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCACTTCCCTTCCCTGTC 330
QY 297 CTTCTCTCAATGGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGCCACTTCCCTTCCCTGTC 356
|||||
Db 331 ATTCTCTCAATGGTGTATGCCAGCTGTGCCGACACCTGTGCTGCTGCTGCTGCTGCTGCTGCT 390
QY 357 CCCAGTCCCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAAGTGTGT 416
|||||
Db 391 CCCAGTCCCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGCTGCTGCTGCTGCTGCT 450
QY 417 GCACGAGGCTGGGGAGTCTTGGACACCTGCATGTCTGGACCCCGAGCCAGGCGCTG 476
|||||
Db 451 GCACGAGGCTGGGGAGTCTTGGACACCTGCATGTCTGGACCCCGAGCCAGGCGCTG 510
QY 477 GTTGTGACGCTGGGGCAGGCGCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAG 533
|||||
Db 511 GTTGTGACGCTGGGGCAGGCGCGCTGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 567

RESULT 11
BM921531
LOCUS BM921531 1006 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6708025 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753009
5', mRNA sequence.
ACCESSION BM921531
VERSION BM921531.1 GI:19371910
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1006)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12788 row: b column: 18
High quality sequence stop: 593.
Location/Qualifiers
1. .1006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5753009"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:

FEATURES
source

pcMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 148 a 368 c 317 g 169 t 4 others
ORIGIN
Query Match 24.7%; Score 421.2; DB 14; Length 1006;
Best Local Similarity 77.5%; Pred. No. 2.4e-103;
Matches 585; Conservative 0; Mismatches 161; Indels 9; Gaps 6;
QY 256 GCAGCCCACTGATCCATCTCTGGCCACTTCCCTTCTCTCTCAATGGTGTGTG 315
|||||
Db 1 GCACACCGAAGACCCACTCTCGGCTTCTCCCTCTCTCTCAAAAGTGGCTA 60
QY 316 CCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTCCCCACAGGGG 375
|||||
Db 61 CCAGCTGTGCCGGACACCAATGTACCTGCCCCCTGGCCACTCCCCGATGCCCGTGGGAG 120
QY 376 TACCCCTGGTGTGGATGGCTGTGGCTGTCTTAAAGTGTGTGACGGAGGCTGGGGGAGT 435
|||||
Db 121 TACCCCTGGTGTGGATGGCTGTGGCTGTCTCCGGGTATGTGACGGCGGCTGGGGGAGC 180
QY 436 CTGCGACCACTGCATGTCTGCGACCCAGCCAGGCGCTGGTGTGTGAGCCCTGGGCGAG 495
|||||
Db 181 CCTGCGACCAACTCCACGCTCTGCGACGCCAGCCAGGCGCTGTGTGCCAGCCCGGGGCGAG 240
QY 496 GCCCTGGCGGCCATGGGGCTGTGTCTTCTTGGATGAGGATGACGGTAGCTGTGAGGTGA 555
|||||
Db 241 GACCCGGTGGACGGGGGCGCTGTGCTTCTTGGCAGAGGACGACAGCAGCTGTGAGGTGA 300
QY 556 ATGCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTCAGGGTCTCTGCGCGCT 615
|||||
Db 301 ACGCGCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCT 360
QY 616 GTGATGACGGTGGCTTCACTGCTGCTGCCGCTGTGTCAGTGAAGTGTGCGGCTGCCAGCT 675
|||||
Db 361 GCGAGGACGGCGGCTTCACTGCTGCTGCCGCTGTGTCAGCGAGATGTGCGGCTGCCAGCT 420
QY 676 GGGACTGCCCCACGCCCCAAGAGAATACAGGTGCCAGGAAAGTGTGCCCGAGTGGGTAT 735
|||||
Db 421 GGGACTGCCCCACGCCCCAGGGGTCGAGGTCTCTGGCAAGTGTGCTGCTGCTGCTGCTGCT 480
QY 736 GTGACCAAGGAGTGACACCGGGGATCCAGCGCTCCACGGCGCAAGGACACCAACTTTCTG 795
|||||
Db 481 GCGGCCAAGGAGGGGACTGGGACCCAGCCCCCTTCCAGCCCAAGGACCCAGTTTCTG 540
QY 796 CCCTTGTCACTCCTGCTGCTGATGCTCTTGTCTTCCAAATTTGGAGCAC-AGCCTGGGGC 854
|||||
Db 541 GCCTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCCCAGAAATGAGCACGCGCTGGGGA 600
QY 855 CCCTGCTCAACCAACC- -TGTGGCTGGGCATAG-CCACCCAGTGTCCAA--CCAGAACC 909
|||||
Db 601 CCCTGCTCGACCAACCCCTGTGGCTGNGCATGGCCACCCCGGCTGTCCAAACCAAGAACCN 660
QY 910 GATTCTGCCAACTGGAGATCCAACG-CCGCCTGTGTGTGCCAGACCCCTGCCTGGCAGCC 968
|||||
Db 661 GCTTCTGCCGACTGGAGACCCAGCGCCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 969 A--GGAGCCACAGCTCATGGAACAGTGTCTTTCTAA 1001
|||||
Db 721 AGGGGGTTCGACGTCCCCCAAAACAGTGCCTTTCTAA 755

RESULT 12
BQ073722
LOCUS BQ073722 1073 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_7046577 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:5806602

Db 2 GGTGTCGCCGGGTATGTCACGGCGGCTGGGGAGCCCTGGACCAACTCCACGTCGTC 61

Qy 459 GACCCAGCCAGGCGCTGTTGTCAGCCTGGGGCAGGCCCTGGCGCCATGGGGCTGTG 518

Db 62 GACGCAGCCAGGCGCTGCTGCCAGCCCGGGCAGGACCCGGTGGCCGGGGGCCCTG 121

Qy 519 TGTCTCT-----TGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGCAGGTACC 570

Db 122 TGCCTCTGTAAGCAGTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGGCCGCTGTATC 181

Qy 571 TGGATGAGAGACCTTTAAACCCCAATTGCAGGGTCCTGTGCCGCTGTGATGACGGTGGCT 630

Db 182 GGGAGGGGAGACCTTCCAGCCCCCACTGCAGCATCCGGCTGCCGTGCGAGGACGGCGGCT 241

Qy 631 TCACCTGCCTGCCGCTGTGCAGTGAGGATGTGCGGGTGCCAGCTGGGACTGCCACGCC 690

Db 242 TCACCTGCGTGCCGCTGTGCAGCAGGATGTGCGGCTGCCAGCTGGGACTGCCCCACC 301

Qy 691 CCAAGAGAATACAGGTGCCAGGAAAGTGTGCCCCGAGTGGGTATGTGACCAGGGAGTGA 750

Db 302 CCAGAGGGTTCGAGGTCTCTGGCAAGTGTGCCCTGAGTGGTGTGCGGCCAAGGAGGG 361

Qy 751 CACGGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCCCTTGTCACTCCTG 810

Db 362 GACTGGGGACCCAGCCCCCTTCCAGCCCAAGGACCCCACTTCTGGCCTTGTCTCTCCC 421

Qy 811 CCTCTGCTGATGCTCCTTGTCCAAATTGGAGCACAGCCCTGGGGCCCCCTGCTCAACCACT 870

Db 422 TGCCCCCTGGTGTCCCCCTGCCAGAATGGACACGGCCTGGGGACCCCTGCTCGACCACT 481

Qy 871 GTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAACCGATTCTGCCAACTGGAGATCC 930

Db 482 GTGGGCTGGGCATAGCCACCCGGGTGTCCAACCAAGAACCGATTCTGCCGACTGGAGACC 541

Qy 931 AACGGCGCCTGTGTCTGCCAGACCCCTGCCTGGCAGCCAGGACCCACAGCTCATGGAACA 990

Db 542 AGCGCGCCTGTGCCTGTCTCAGGCGCTGCCACCCCTCCAGGGGTGCGAGTCCACAAAACA 601

Qy 991 GTGCTTTCTA-AGGCCAACTGGGATGCGGATACAGGGCCTGCCATCCTCAGCAAAATGAC 1049

Db 602 GTGCTTCTAGAGCCGGGCTGGGAATGGGACACGGTGTCCACCATCCCCAGCTGGTGGC 661

Qy 1050 CCTAGACCAAGGCCCTGG----ACTGCTGGTAGATGCTCTTCTCCATGCTCTTGGCTGCA 1105

Db 662 CCGTGCCTGGGCCCTGGGGCTGATGGAACGATGGTCCGTGCCAGGGCCCTTGGTTGCA 721

Qy 1106 GTTAACTGTCCTGCTTGGATTCACTGTGTAGAGC 1139

Db 722 GGCAACACTTTAGCTTGGGTCCACCATGCAGAAC 755

RESULT 14
BI826781

LOCUS 800 bp mRNA linear EST 04-OCT-2001

DEFINITION 603077268F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169065 5', mRNA sequence.

ACCESSION BI826781

VERSION BI826781.1 GI:15938331

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 800)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM11420 row: g column: 18
High quality sequence stop: 788.

FEATURES

Location/Qualifiers

1..800

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5169065"

/clone_lib="NIH_MGC_119"

/tissue_type="medulla"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

BASE COUNT 115 a 284 c 264 g 136 t 1 others

ORIGIN

Query Match 24.0%; Score 410.2; DB 13; Length 800;
Best Local Similarity 77.5%; Pred. No. 2.2e-100;
Matches 572; Conservative 0; Mismatches 159; Indels 7; Gaps 6;

Qy 269 CCATCTTCTGGCCACTTCCCTTCTCTGCTTCTCTCAATGGTGTGCCCCAGCTGTGCCG 328

Db 14 CCACCTCCTGGCCTTCTCCCTCCTCTGCTCCTCTCAAAAGTGGTACCCAGCTGTGCC 73

Qy 329 GACACCCCTGTACCTGTCTTGGACACACCCCACTGCCCCACAGGGGGTACCCCTGGTCT 388

Db 74 GACACCATGTACCTGCCCCCTGGCCACCTCCCCGATGCCCCGTGGGAGTACCCCTGGTCT 133

Qy 389 GGATGGCTGTGGCTGTGTAAGTGTGTGCACGGAGGTGGGGGAGTCTCGGACCACT 448

Db 134 GGATGGCTGTGGCTGTGCCGGGTATGTGCACGGCGGTGGGGAGCCCTGCGACCAACT 193

Qy 449 GCATGTCTGCGACCCCAAGCCCTGGTTTGTGAGCTGGGGCAGGCCCTGGCGGCCA 508

Db 194 CCAGCTCTGCGACGCCAGCCA-GGCCTGGTCTGCCAGCCCGGGGAGGACCCGGTGCCG 252

Qy 509 TGGGCTGTGTGCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGAGGTA 568

Db 253 GGGGCCCCCTGTGCTCTTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGGCCGCTGTA 312

Qy 569 CCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTCTGTCGGCTGTGATGACGGTGG 628

Db 313 TCGGAAGGGGAGACCTTCCAGCCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGG 372

Qy 629 CTTACCTGCCTGCCGCTGTGCAGTGAGGATGTGCGGGTGCCTGCTGCTGCTGCTGCTGCTG 688

Db 373 CTTACCTGCCTGCCGCTGTGCAGGAGGATGTGCGGGCTGCCAGCTGGGACTGCCCCCA 432

Qy 689 CCCCAGAGAAATACAGGTGCCAGGAAAGTGTGCCCCGAGTGGGTATGTGACCAGGGAGT 748

Db 433 CCCCAGGAGGTTCAGGTCTTGGGCAAGTGTGCCCTGAGTGGGTGTGCG-GCAAGGAGG 491

Qy 749 GACACCGCGCATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCCCTTGTCACTCC 808

Db 492 GGGACTGGGGACCCAGCCCCCTTCCAGCCCAAGGACCCCACTTCTTGGCCTTGTCTCTTC 551

Qy 809 TGCCTCTGTGATGCTCCTTGTCCAAATTGGAGCACAGCCTGGGGCCCCCTGCTCAACCCAC 868

Db 552 -CCTGCCCCCTGGTGTCCCCCTGCCAGAATGGAGNCACGGC--TGGGACCCCTGCTCGACCCAC 608

Qy 869 CTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAACCAGAACCCGATTCTGCCAACTGGAGAT 928

Db 609 CTGTGGGCTGGGCATGGCCACCCCGGGGTGTCCAAACCAGAACCG-TTCTGCCGACTGGAGAC 667

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:59:01 ; Search time 486.369 Seconds
(without alignments)
10471.464 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708
Perfect score: 175
Sequence: 1 AGTCCAGGAACCTTGAGCTTT.....GCCTAGATAAACACCCAAA 175

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	171.8	98.2	137964	2	AC126895	AC126895 Rattus no
2	171.8	98.2	226303	2	AC095418	AC095418 Rattus no
3	156.6	89.5	1741	10	AF259981	AF259981 Rattus no
4	126.2	72.1	1734	6	AR210324	AR210324 Sequence
5	126.2	72.1	1734	6	AR210325	AR210325 Sequence
6	126.2	72.1	1734	10	AF100778	AF100778 Mus muscu
7	126.2	72.1	61072	10	AL731698	AL731698 Mouse DNA
8	126.2	72.1	216757	2	AL669906	AL669906 Mus muscu
9	121.4	69.4	1739	10	AF126063	AF126063 Mus muscu
10	36.4	20.8	208219	2	AC130013	AC130013 Rattus no
11	35	20.0	112855	2	AP003608	AP003608 Oryza sat
12	34.8	19.9	143565	2	AC023195	AC023195 Homo sapi
13	34.8	19.9	190964	10	AL671848	AL671848 Mouse DNA
14	34.8	19.9	233269	2	AC094291	AC094291 Rattus no
15	34.4	19.7	144233	2	AC068103	AC068103 Homo sapi
16	34.4	19.7	161900	9	AC096729	AC096729 Homo sapi
17	33.6	19.2	238910	2	AC122851	AC122851 Mus muscu
18	33.4	19.1	167633	2	AC129096	AC129096 Papio cyn
19	33	18.9	171564	2	AC113225	AC113225 Rattus no
20	33	18.9	173085	2	AC112105	AC112105 Rattus no
21	33	18.9	181719	2	AL662863	AL662863 Mus muscu
22	33	18.9	195339	10	AL596181	AL596181 Mouse DNA
23	33	18.9	210487	9	AL445199	AL445199 Human DNA
24	32.8	18.7	145832	9	AC090736	AC090736 Homo sapi
25	32.8	18.7	202317	9	AC020603	AC020603 Homo sapi
26	32.8	18.7	203226	2	AC109624	AC109624 Mus muscu
27	32.8	18.7	213438	2	AL844489	AL844489 Mus muscu
28	32.4	18.5	65992	2	AC107972	AC107972 Homo sapi
29	32.4	18.5	160199	9	AF391285	AF391285 Homo sapi
30	32.4	18.5	182522	9	AC015689	AC015689 Homo sapi
31	32.2	18.4	119531	2	AP003879	AP003879 Oryza sat
32	32.2	18.4	150428	2	AP004591	AP004591 Oryza sat
33	32	18.3	173755	2	AC121976	AC121976 Mus muscu
34	32	18.3	185536	2	AC122340	AC122340 Mus muscu
35	32	18.3	187490	2	AC048378	AC048378 Homo sapi
36	32	18.3	192404	2	AC128523	AC128523 Rattus no
37	32	18.3	202634	2	AC114771	AC114771 Homo sapi
38	32	18.3	212421	2	AC109591	AC109591 Homo sapi
39	32	18.3	221647	10	AL591003	AL591003 Mouse DNA
40	32	18.3	227103	2	AC122470	AC122470 Mus muscu
41	31.8	18.2	110000	2	AC073702_2	Continuation (3 of
42	31.8	18.2	132278	9	AC026784	AC026784 Homo sapi
43	31.8	18.2	147817	8	AP003267	AP003267 Oryza sat
44	31.8	18.2	151269	8	AC016781	AC016781 Genomic S
45	31.8	18.2	151491	2	AC084764	AC084764 Oryza sat

ALIGNMENTS

RESULT 1
AC126895

LOCUS
DEFINITION

AC126895 137964 bp DNA linear HTG 24-JUL-2002
Rattus norvegicus clone CH230-301E4, *** SEQUENCING IN PROGRESS
***, 49 unordered pieces.

ACCESSION

AC126895.1 GI:21724040

VERSION
HTG; HTGS_PHASE1.

KEYWORDS
Rattus norvegicus.

SOURCE

ORGANISM
Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE
1 (bases 1 to 137964)

AUTHORS

Muzny,D.M., Adams,C., Ali-Osman,F.R., Allen,C.,

Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Slissov, I.,
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlarczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 137964)
Worley, K.C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 137964)
Worley, K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZHG
Center clone name: CH230-301E4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89474 bases at least Q40
Consensus quality: 93422 bases at least Q30
Consensus quality: 96506 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 1056: contig of 1056 bp in length
* 1057: gap of unknown length
* 1157: contig of 1179 bp in length
* 2335: contig of 1179 bp in length
* 2336: gap of unknown length
* 2435: gap of unknown length
* 3442: contig of 1007 bp in length
* 3443: gap of unknown length
* 3542: gap of unknown length
* 3543: contig of 1539 bp in length
* 5081: contig of 1539 bp in length
* 5181: gap of unknown length
* 5182: contig of 1243 bp in length
* 6424: contig of 1243 bp in length
* 6524: gap of unknown length
* 7814: contig of 1290 bp in length
* 7914: gap of unknown length
* 9153: contig of 1239 bp in length
* 9154: gap of unknown length
* 9254: contig of 1192 bp in length
* 10445: gap of unknown length
* 10545: contig of 1192 bp in length
* 10546: gap of unknown length
* 11771: contig of 1226 bp in length
* 11871: gap of unknown length
* 13240: contig of 1369 bp in length
* 13340: gap of unknown length
* 13341: contig of 1019 bp in length
* 14359: gap of unknown length
* 14360: gap of unknown length
* 14459: gap of unknown length
* 14651: contig of 1192 bp in length
* 15651: gap of unknown length
* 15751: gap of unknown length
* 17494: contig of 1743 bp in length
* 17495: gap of unknown length
* 17594: gap of unknown length
* 17595: contig of 1085 bp in length
* 18679: gap of unknown length
* 18779: gap of unknown length
* 20681: contig of 1902 bp in length
* 20781: gap of unknown length
* 20782: contig of 1337 bp in length
* 22118: gap of unknown length
* 22218: gap of unknown length
* 23578: contig of 1360 bp in length
* 23579: gap of unknown length
* 23678: gap of unknown length
* 23679: contig of 1745 bp in length
* 25423: gap of unknown length
* 25424: gap of unknown length
* 25523: gap of unknown length
* 25524: contig of 2285 bp in length
* 27808: gap of unknown length
* 27809: gap of unknown length
* 27909: contig of 2364 bp in length
* 30272: gap of unknown length
* 30273: gap of unknown length
* 30373: contig of 2719 bp in length
* 33091: gap of unknown length
* 33092: gap of unknown length
* 33191: gap of unknown length
* 33192: contig of 2586 bp in length
* 35777: gap of unknown length
* 35778: gap of unknown length
* 35877: gap of unknown length
* 35878: contig of 1031 bp in length
* 36908: gap of unknown length
* 37008: gap of unknown length
* 37009: contig of 1922 bp in length
* 38930: gap of unknown length
* 38931: gap of unknown length
* 39030: gap of unknown length
* 39031: contig of 2180 bp in length
* 41210: gap of unknown length
* 41211: gap of unknown length
* 41310: gap of unknown length
* 41311: contig of 1627 bp in length
* 42937: gap of unknown length
* 42938: gap of unknown length
* 43037: gap of unknown length
* 43038: contig of 3083 bp in length
* 46120: gap of unknown length
* 46121: gap of unknown length
* 46220: gap of unknown length
* 46221: contig of 1997 bp in length
* 48217: gap of unknown length
* 48218: gap of unknown length
* 48317: gap of unknown length
* 50433: contig of 2116 bp in length
* 50434: gap of unknown length
* 50533: gap of unknown length
* 52880: contig of 2347 bp in length
* 52881: gap of unknown length
* 52980: gap of unknown length
* 55194: contig of 2214 bp in length
* 55195: gap of unknown length
* 55294: gap of unknown length
* 55295: contig of 2188 bp in length
* 57482: gap of unknown length
* 57483: gap of unknown length
* 57582: gap of unknown length
* 60781: contig of 3199 bp in length
* 60782: gap of unknown length
* 60881: gap of unknown length
* 62599: contig of 1718 bp in length

*	62600	62699: gap of unknown length	*	62699: gap of unknown length
*	62700	65203: contig of 2504 bp in length	*	65203: contig of 2504 bp in length
*	65204	65303: gap of unknown length	*	65303: gap of unknown length
*	65304	66914: contig of 1611 bp in length	*	66914: contig of 1611 bp in length
*	66915	67014: gap of unknown length	*	67014: gap of unknown length
*	67015	70839: contig of 3825 bp in length	*	70839: contig of 3825 bp in length
*	70840	70939: gap of unknown length	*	70939: gap of unknown length
*	70940	75139: contig of 4200 bp in length	*	75139: contig of 4200 bp in length
*	75140	75239: gap of unknown length	*	75239: gap of unknown length
*	75240	78296: contig of 3057 bp in length	*	78296: contig of 3057 bp in length
*	78297	78396: gap of unknown length	*	78396: gap of unknown length
*	78397	83138: contig of 4742 bp in length	*	83138: contig of 4742 bp in length
*	83139	83238: gap of unknown length	*	83238: gap of unknown length
*	83239	88204: contig of 4966 bp in length	*	88204: contig of 4966 bp in length
*	88205	88304: gap of unknown length	*	88304: gap of unknown length
*	88305	92238: contig of 3934 bp in length	*	92238: contig of 3934 bp in length
*	92239	92338: gap of unknown length	*	92338: gap of unknown length
*	92339	97339: contig of 5001 bp in length	*	97339: contig of 5001 bp in length
*	97340	97439: gap of unknown length	*	97439: gap of unknown length
*	97440	103534: contig of 6095 bp in length	*	103534: contig of 6095 bp in length
*	103535	103634: gap of unknown length	*	103634: gap of unknown length
*	103635	107080: contig of 3446 bp in length	*	107080: contig of 3446 bp in length
*	107081	107180: gap of unknown length	*	107180: gap of unknown length
*	107181	115201: contig of 8021 bp in length	*	115201: contig of 8021 bp in length
*	115202	115301: gap of unknown length	*	115301: gap of unknown length
*	115302	121237: contig of 5936 bp in length	*	121237: contig of 5936 bp in length
*	121238	121337: gap of unknown length	*	121337: gap of unknown length
*	121338	127910: contig of 6573 bp in length	*	127910: contig of 6573 bp in length
*	127911	128010: gap of unknown length	*	128010: gap of unknown length
*	128011	137964: contig of 9954 bp in length.	*	137964: contig of 9954 bp in length.
FEATURES				
Source				
1..137964				
/organism="Rattus norvegicus"				
/db_xref="taxon:10116"				
/clone="CH230-301E4"				
BASE COUNT	31831 a	34027 c	35229 g	31979 t 4898 others
ORIGIN				
Query Match	98.2%; Score 171.8; DB 2; Length 137964;			
Best Local Similarity	98.9%; Pred. No. 6.9e-50;			
Matches 173; Conservative	0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	AGTCCAGGAACCTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAC	60	
Db	44819	AGTCCAGGAACCTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAC	44878	
QY	61	AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGAAAGACAAGG	120	
Db	44879	AGGAAGGCTCCACACCTCTGGTAGGCCAGGGCCTTCTCTTCAGCATGAGAAAGACAAGG	44938	
QY	121	GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCTAGAATAAACACCCAAA	175	
Db	44939	GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCCAGCCTAGAATAAACACCCAAA	44993	
RESULT 2				
AC095418				
LOCUS	226303 bp DNA linear HTG 11-JUL-2002			
DEFINITION	Rattus norvegicus clone CH230-7C10, *** SEQUENCING IN PROGRESS ***, 51 unordered pieces.			
ACCESSION	AC095418			
VERSION	AC095418.3 GI:21717893			
KEYWORDS	HTG; HTGS_PHASE1.			
SOURCE	Norway rat.			
ORGANISM	Rattus norvegicus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
Rattus.				
REFERENCE	1 (bases 1 to 226303)			
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,			

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.	TITLE	JOURNAL	
Direct Submission			
Unpublished	REFERENCE	JOURNAL	
2 (bases 1 to 226303)			
Worley,K.C.	TITLE	JOURNAL	
Direct Submission			
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	REFERENCE	JOURNAL	
3 (bases 1 to 226303)			
Worley,K.C.	TITLE	JOURNAL	
Direct Submission			
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	COMMENT		
On Jul 10, 2002 this sequence version replaced gi:17941885.			
Center: Baylor College of Medicine			
Center code: BCM			
Web site: http://www.hgsc.bcm.tmc.edu/			
Contact: hgsc-help@bcm.tmc.edu			
----- Project Information			
Center project name: GCDF			
Center clone name: CH230-7C10			
----- Summary Statistics			
Sequencing vector: Plasmid;			
Chemistry: Dye-terminator Big Dye; 100% of reads			
Assembly program: Phrap; version 0.990329			
Consensus quality: 185950 bases at least Q40			
Consensus quality: 190362 bases at least Q30			
Consensus quality: 193076 bases at least Q20			

* NOTE: Estimated insert size may differ from sequence length			
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).			
* NOTE: This is a 'working draft' sequence. It currently			
* consists of 51 contigs. The true order of the pieces			
* is not known and their order in this sequence record is			
* arbitrary. Gaps between the contigs are represented as			

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1051: contig of 1051 bp in length
* 1052 1151: gap of unknown length
* 1152 2286: contig of 1135 bp in length
* 2287 2386: gap of unknown length
* 2387 3433: contig of 1047 bp in length
* 3434 3533: gap of unknown length
* 3534 4614: contig of 1081 bp in length
* 4615 4714: gap of unknown length
* 4715 6581: contig of 1867 bp in length
* 6582 6681: gap of unknown length
* 6682 7773: contig of 1092 bp in length
* 7774 7873: gap of unknown length
* 7874 9530: contig of 1657 bp in length
* 9531 9630: gap of unknown length
* 9631 10661: contig of 1031 bp in length
* 10662 10761: gap of unknown length
* 10762 11958: contig of 1197 bp in length
* 11959 12058: gap of unknown length
* 12059 13140: contig of 1082 bp in length
* 13141 13240: gap of unknown length
* 13241 14777: contig of 1537 bp in length
* 14778 14877: gap of unknown length
* 14878 16063: contig of 1186 bp in length
* 16064 16163: gap of unknown length
* 16164 17585: contig of 1422 bp in length
* 17586 17685: gap of unknown length
* 17686 18916: contig of 1231 bp in length
* 18917 19016: gap of unknown length
* 19017 20305: contig of 1289 bp in length
* 20306 20405: gap of unknown length
* 20406 21537: contig of 1132 bp in length
* 21538 21637: gap of unknown length
* 21638 23703: contig of 2066 bp in length
* 23704 23803: gap of unknown length
* 23804 26306: contig of 2503 bp in length
* 26307 26406: gap of unknown length
* 26407 28431: contig of 2025 bp in length
* 28432 28531: gap of unknown length
* 28532 30615: contig of 2084 bp in length
* 30616 30715: gap of unknown length
* 30716 33215: contig of 2500 bp in length
* 33216 33315: gap of unknown length
* 33316 35357: contig of 2042 bp in length
* 35358 35457: gap of unknown length
* 35458 38012: contig of 2555 bp in length
* 38013 38112: gap of unknown length
* 38113 40338: contig of 2226 bp in length
* 40339 40438: gap of unknown length
* 40439 44051: contig of 3613 bp in length
* 44052 44151: gap of unknown length
* 44152 46294: contig of 2143 bp in length
* 46295 46394: gap of unknown length
* 46395 48845: contig of 2451 bp in length
* 48846 48945: gap of unknown length
* 48946 51753: contig of 2808 bp in length
* 51754 51853: gap of unknown length
* 51854 55915: contig of 4062 bp in length
* 55916 56015: gap of unknown length
* 56016 60468: contig of 4453 bp in length
* 60469 60568: gap of unknown length
* 60569 63640: contig of 3072 bp in length
* 63641 63740: gap of unknown length
* 63741 67888: contig of 4148 bp in length
* 67889 67988: gap of unknown length
* 67989 73014: contig of 5026 bp in length
* 73015 73114: gap of unknown length
* 73115 77313: contig of 4199 bp in length
* 77314 77413: gap of unknown length
* 77414 83786: contig of 6373 bp in length

* 83787 83886: gap of unknown length
* 83887 89590: contig of 5704 bp in length
* 89591 89690: gap of unknown length
* 89691 93846: contig of 4156 bp in length
* 93847 93946: gap of unknown length
* 93947 98126: contig of 4180 bp in length
* 98127 98226: gap of unknown length
* 98227 104258: contig of 6032 bp in length
* 104259 104358: gap of unknown length
* 104359 110994: contig of 6636 bp in length
* 110995 111094: gap of unknown length
* 111095 119838: contig of 8744 bp in length
* 119839 119938: gap of unknown length
* 119939 126339: contig of 6401 bp in length
* 126340 126439: gap of unknown length
* 126440 134888: contig of 8449 bp in length
* 134889 134988: gap of unknown length
* 134989 142709: contig of 7721 bp in length
* 142710 142809: gap of unknown length
* 142810 149818: contig of 7009 bp in length
* 149819 149918: gap of unknown length
* 149919 158719: contig of 8801 bp in length
* 158720 158819: gap of unknown length
* 158820 169868: contig of 11049 bp in length
* 169869 169968: gap of unknown length
* 169969 180390: contig of 10422 bp in length
* 180391 180490: gap of unknown length
* 180491 191272: contig of 10782 bp in length
* 191273 191372: gap of unknown length
* 191373 209667: contig of 18295 bp in length
* 209668 209767: gap of unknown length
* 209768 226303: contig of 16536 bp in length.

FEATURES

source

1..226303
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-7C10"

BASE COUNT 57875 a 53353 c 52917 g 57041 t 5117 others

Query Match 98.2%; Score 171.8; DB 2; Length 226303;
Best Local Similarity 98.9%; Pred. No. 7.5e-50;
Matches 173; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 113012 AGTCCAGGAACCTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 113071
|||||
QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCCCTTCTCTTCCAGCATGAGAAAGACAAGG 120
|||||
Db 113072 AGGAAGGCTCCACACCTCTGGTAGGCCAGGCCCTTCTCTTCCAGCATGAGAAAGACAAGG 113131
|||||
QY 121 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175
|||||
Db 113132 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 113186
|||||

RESULT 3

AF259981

LOCUS

DEFINITION

AF259981.1 GI:7739780

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus.

Rattus norvegicus

Eukaryota; Metazoa;

Mammalia; Eutheria;

Rodentia; Sciurognathi;

Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 1741)

Zhang, R., Averboukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P.J.,

Coffey, R.J., Pardee, A.B. and Liang, P.

Identification of rCop-1, a new member of the CCN protein family,

AF259981 1741 bp mRNA linear ROD 09-MAY-2000
Rattus norvegicus CCN family protein COP-1 (Cop-1) mRNA, complete

as a negative regulator for cell transformation
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)
98414629

REFERENCE 2 (bases 1 to 1741)
AUTHORS Liang,P.
TITLE Direct Submission

JOURNAL Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer
Center, 649 MRB II, Nashville, TN 37232, USA
FEATURES Location/Qualifiers

source 1..1741
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
gene 1..1741
/gene="Cop-1"
CDS 262..1014
/gene="Cop-1"
/note="secreted protein"
/codon_start=1
/product="CCN family protein COP-1"
/protein_id="AAF69011.1"
/db_xref="GI:7739781"
/translation="MRGSLIRLLATSFLLCSMVCAQLCRTPCTCPWTPPQCPOGVP
LVLDGCGCKVCARRLTESCEHLHVCEPSQGLVCQPGAGPGGHGAVCLLDEDDGCEV
NGRRYLDGETFKPNCRVLCRDDGFTCLPLCSEDVTLPSWDCPRPKRIQVPGKCCPE
WVCDQGVTPAIQRSAAQHQLSALVTPASADAPWPNWSTANGPCSTTCGLGIATRVSN
QNRFCQLEIQRRRLCLPRPCLAAARSHSSWNSAF"

BASE COUNT 386 a 491 c 480 g 384 t
ORIGIN
Query Match 89.5%; Score 156.6; DB 10; Length 1741;
Best Local Similarity 97.1%; Pred. No. 8.8e-45;
Matches 170; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1547 AGTCCAGGAGCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 1606
QY 61 AGGAAGCTCCACACCTCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGAAAGACAAGG 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1607 AGGAAGCTCCACACCTCTAACA-GCCAGGCCCTTCTCTTCAGCATGAGAAAGACAAGG 1665
QY 121 GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1666 GACAGCAGAGTACTCTCGTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 1720

RESULT 4
AR210324
LOCUS AR210324 1734 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 17 from patent US 6387657.
ACCESSION AR210324
VERSION AR210324.1 GI:21512525
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 17 14-MAY-2002;
FEATURES Location/Qualifiers
source 1..1734
/organism="unknown"
BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN

Query Match 72.1%; Score 126.2; DB 6; Length 1734;
Best Local Similarity 86.3%; Pred. No. 7e-34;
Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60

Db 1536 AGTCCAGGAACCTTGACCTTTGTATTGTAAAAATACACATCTCTTAAATGCTCACAAGC 1595
QY 61 AGGAAGCTCCACACCTCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGAAAGACAAGG 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1596 AAG-AGGCTCCACACTTCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGAGACAAGG 1654
QY 121 GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1655 AACAGTAGAGTACCCTCCTCTGTGGAGGACTGGCCGGTCTGGATAAACACCCAAA 1709

RESULT 5
AR210325/c
LOCUS AR210325 1734 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 18 from patent US 6387657.
ACCESSION AR210325
VERSION AR210325.1 GI:21512526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 18 14-MAY-2002;
FEATURES Location/Qualifiers
source 1..1734
/organism="unknown"

BASE COUNT 393 a 495 c 491 g 355 t
ORIGIN
Query Match 72.1%; Score 126.2; DB 6; Length 1734;
Best Local Similarity 86.3%; Pred. No. 7e-34;
Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;
QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 199 AGTCCAGGAACCTTGACCTTTGTATTGTAAAAATACACATCTCTTAAATGCTCACAAGC 140
QY 61 AGGAAGCTCCACACCTCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGAAAGACAAGG 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 139 AAG-AGGCTCCACACTTCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGAGACAAGG 81
QY 121 GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 80 AACAGTAGAGTACCCTCCTCTGTGGAGGACTGGCCGGTCTGGATAAACACCCAAA 26

RESULT 6
AF100778
LOCUS AF100778 1734 bp mRNA linear ROD 17-DEC-1998
DEFINITION Mus musculus connective tissue growth factor related protein WISP-2
(Wisp2) mRNA, complete cds.
ACCESSION AF100778
VERSION AF100778.1 GI:4028578
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1734)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A.,
Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C.,
Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D.,
Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.
TITLE WISP genes are members of the connective tissue growth factor
family that are up-regulated in wnt-1-transformed cells and
aberrantly expressed in human colon tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)
MEDLINE 99061933
PUBMED 9843955

REFERENCE 2 (bases 1 to 1734)
AUTHORS Pennica,D.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA

FEATURES
source
1..1734
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="C57MG"
/cell_type="epithelial"
/tissue_type="mammary"
/note="transformed by Wnt-1"
1..1734
/gene="Wisp2"
257..1012
/gene="Wisp2"
/codon_start=1
/product="connective tissue growth factor related protein WISP-2"
/protein_id="AAC96320.1"
/db_xref="GI:4028579"
/translation="MRGNPLIHLAISFLCILSMVYSQLCPAPCAPWTPPPQPPGVP
LVLDGCGCRVCARRLGESCDHLHVCDPVSGLVCPGAGPSGRGAVCLFEEDDGSCEV
NGRRYLDGETFKPNCRVLCRDDGGFTCLPLCSEDEVLPVSWDCPRPRRIQVPGRCPE
WVCDQAVMQPAIQPNCSAOGHQLSALVTPASADGPCPNWSTAWPCSTTCGLGIATRV
NQNRFCQLEIQRLCLSRPCLASRSHGSWNSAF"

BASE COUNT 355 a 491 c 495 g 393 t

ORIGIN
Query Match 72.1%; Score 126.2; DB 10; Length 1734;
Best Local Similarity 86.3%; Pred. No. 7e-34;
Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAC 60
|||||
Db 1536 AGTCCAGGAACCTTGACCTTTGTATTGTAAAATAACACATCTCTTAATGCTCACAAGC 1595
|||||

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAGG 120
| | | | |
Db 1596 AAG-AGGCTCCACACTTCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAGACAAGG 1654
| | | | |

QY 121 GACAGCAGAGTACTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175
|||||

Db 1655 AACAGTAGAGTACCTCTCTGGAGGACTGGCCGGTCTGGAATAAACACCCAAA 1709
|||||

RESULT 7
AL731698
LOCUS Mouse DNA sequence from clone RP23-161B3 on chromosome 2, complete sequence.
DEFINITION AL731698 61072 bp DNA linear ROD 24-MAY-2002
ACCESSION AL731698
VERSION AL731698.10 GI:21214309
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 61072)
AUTHORS Wallis,J.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On May 25, 2002 this sequence version replaced gi:21213601.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-161B3 is from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe36.

FEATURES
source
1..61072
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-161B3"
/clone_lib="RPCI-23"

BASE COUNT 15405 a 15368 c 15033 g 15266 t

ORIGIN
Query Match 72.1%; Score 126.2; DB 10; Length 61072;
Best Local Similarity 86.3%; Pred. No. 1.3e-33;
Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAC 60
|||||
Db 17748 AGTCCAGGAACCTTGACCTTTGTATTGTAAAATAACACATCTCTTAATGCTCACAAGC 17807
|||||

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAGG 120
| | | | |
Db 17808 AAG-AGGCTCCACACTTCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAGACAAGG 17866
| | | | |

QY 121 GACAGCAGAGTACTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175
|||||

Db 17867 AACAGTAGAGTACCTCTCTGGAGGACTGGCCGGTCTGGAATAAACACCCAAA 17921
|||||

RESULT 8
AL669906
LOCUS Mus musculus chromosome 2 clone RP23-217C2, *** SEQUENCING IN PROGRESS ***, 21 unordered pieces.
DEFINITION AL669906 216757 bp DNA linear HTG 24-JUL-2002
ACCESSION AL669906
VERSION AL669906.5 GI:21955520
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 216757)
AUTHORS Sims,S.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On Jul 25, 2002 this sequence version replaced gi:18181793.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BM217C2
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 211865 bases at least Q40
Consensus quality: 213403 bases at least Q30

Consensus quality: 214139 bases at least Q20
Insert size: 214757; sum-of-contigs
Insert size: 234243; 1.8% error; agarose-fp
Quality coverage: 6.29x in Q20 bases; sum-of-contigs Quality
coverage: 5.90x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4000: contig of 4000 bp in length
* 4001 4100: gap of 100 bp
* 4101 15267: contig of 11167 bp in length
* 15268 15367: gap of 100 bp
* 15368 34970: contig of 19603 bp in length
* 34971 35070: gap of 100 bp
* 35071 39671: contig of 4601 bp in length
* 39672 39771: gap of 100 bp
* 39772 46360: contig of 6589 bp in length
* 46361 46460: gap of 100 bp
* 46461 49809: contig of 3349 bp in length
* 49810 49909: gap of 100 bp
* 49910 54509: contig of 4600 bp in length
* 54510 54609: gap of 100 bp
* 54610 65989: contig of 11380 bp in length
* 65990 66089: gap of 100 bp
* 66090 84635: contig of 18546 bp in length
* 84636 84735: gap of 100 bp
* 84736 87399: contig of 2664 bp in length
* 87400 87499: gap of 100 bp
* 87500 98601: contig of 11102 bp in length
* 98602 98701: gap of 100 bp
* 98702 105522: contig of 6821 bp in length
* 105523 105622: gap of 100 bp
* 105623 119773: contig of 14151 bp in length
* 119774 119873: gap of 100 bp
* 119874 134552: contig of 14679 bp in length
* 134553 134652: gap of 100 bp
* 134653 141202: contig of 6550 bp in length
* 141203 141302: gap of 100 bp
* 141303 145808: contig of 4506 bp in length
* 145809 145908: gap of 100 bp
* 145909 151572: contig of 5664 bp in length
* 151573 151672: gap of 100 bp
* 151673 169686: contig of 18014 bp in length
* 169687 169786: gap of 100 bp
* 169787 200095: contig of 30309 bp in length
* 200096 200195: gap of 100 bp
* 200196 213877: contig of 13682 bp in length
* 213878 213977: gap of 100 bp
* 213978 216757: contig of 2780 bp in length.
FEATURES
Source
Location/Qualifiers
1. .216757
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-217C2"
/clone_lib="RPCI-23"
1. .4000
/note="assembly_fragment:03429
clone_end:SP6
vector_side:left"
4101. .15267
/note="assembly_fragment:00814
fragment_chain:1"
15368. .34970
/note="assembly_fragment:00812
fragment_chain:1"
35071. .39671

misc_feature /note="assembly_fragment:03544
fragment_chain:1"
39772. .46360
/note="assembly_fragment:01438
fragment_chain:1"
46461. .49809
/note="assembly_fragment:03410
fragment_chain:2"
49910. .54509
/note="assembly_fragment:00811
fragment_chain:2"
54610. .65989
/note="assembly_fragment:02217.
fragment_chain:2"
66090. .84635
/note="assembly_fragment:03529
fragment_chain:2"
84736. .87399
/note="assembly_fragment:01019
fragment_chain:3"
87500. .98601
/note="assembly_fragment:01170
fragment_chain:3"
98702. .105522
/note="assembly_fragment:02829
fragment_chain:3"
105623. .119773
/note="assembly_fragment:01476
fragment_chain:4"
119874. .134552
/note="assembly_fragment:00505
fragment_chain:4"
134653. .141202
/note="assembly_fragment:01951
fragment_chain:4"
141303. .145808
/note="assembly_fragment:02637
fragment_chain:5"
145909. .151572
/note="assembly_fragment:02352
fragment_chain:5"
151673. .169686
/note="assembly_fragment:02207
fragment_chain:5"
169787. .200095
/note="assembly_fragment:02849
fragment_chain:6"
200196. .213877
/note="assembly_fragment:01052
fragment_chain:6"
213978. .216757
/note="assembly_fragment:01318
clone_end:T7
vector_side:right"
BASE COUNT 56425 a 51573 c 51904 g 54852 t 2003 others
ORIGIN
Query Match 72.1%; Score 126.2; DB 2; Length 216757;
Best Local Similarity 86.3%; Pred. No. 1.7e-33;
Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;
QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 88854 AGTCCAGGAACCTTGACCTTTGTATTGTAAAAATACACATCTCTTAATGCTCACAAAGC 88913
QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGAAAGACAAG 120
| |
Db 88914 AAG-AGGCTCCACACTTCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGAGACAAGG 88972
QY 121 GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCTAGAAATAACACCCAAA 175
||| |
Db 88973 AACAGTAGAGTACCCTCCTCTGAGGACTGGCCCGGCTCTGGAATAAACACCCAAA 89027

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1068: contig of 1068 bp in length
* 1069 1168: gap of unknown length
* 1169 2363: contig of 1195 bp in length
* 2364 2463: gap of unknown length
* 2464 3682: contig of 1219 bp in length
* 3683 3782: gap of unknown length
* 3783 4897: contig of 1115 bp in length
* 4898 4997: gap of unknown length
* 4998 6251: contig of 1253 bp in length
* 6251 6350: gap of unknown length
* 6351 7467: contig of 1117 bp in length
* 7468 7567: gap of unknown length
* 7568 8742: contig of 1175 bp in length
* 8743 8842: gap of unknown length
* 8843 10039: contig of 1197 bp in length
* 10040 10139: gap of unknown length
* 10140 11793: contig of 1654 bp in length
* 11794 11893: gap of unknown length
* 11894 13371: contig of 1478 bp in length
* 13372 13471: gap of unknown length
* 13472 14864: contig of 1393 bp in length
* 14865 14964: gap of unknown length
* 14965 16460: contig of 1496 bp in length
* 16461 16560: gap of unknown length
* 16561 17663: contig of 1103 bp in length
* 17664 17764: contig of 1518 bp in length
* 17764 19281: contig of 1731 bp in length
* 19282 19381: gap of unknown length
* 19382 21112: contig of 1731 bp in length
* 21113 21212: gap of unknown length
* 21213 22833: contig of 1621 bp in length
* 22834 22933: gap of unknown length
* 22934 24214: contig of 1281 bp in length
* 24215 24314: gap of unknown length
* 24315 25527: contig of 1213 bp in length
* 25528 25627: gap of unknown length
* 25628 27220: contig of 1593 bp in length
* 27221 27320: gap of unknown length
* 27321 28926: contig of 1606 bp in length
* 28927 29026: gap of unknown length
* 29027 30363: contig of 1337 bp in length
* 30364 30463: gap of unknown length
* 30464 31665: contig of 1202 bp in length
* 31666 31765: gap of unknown length
* 31766 33349: contig of 1584 bp in length
* 33350 33449: gap of unknown length
* 33450 36317: contig of 2868 bp in length
* 36318 36417: gap of unknown length
* 36418 37778: contig of 1361 bp in length
* 37779 37878: gap of unknown length
* 37879 39851: contig of 1973 bp in length
* 39852 39951: gap of unknown length
* 39952 42802: contig of 2851 bp in length
* 42803 42902: gap of unknown length
* 42903 45682: contig of 2780 bp in length
* 45683 45782: gap of unknown length
* 45783 49268: contig of 3486 bp in length
* 49269 49368: gap of unknown length
* 49369 51428: contig of 2060 bp in length
* 51429 51528: gap of unknown length
* 51529 54571: contig of 3043 bp in length
* 54572 54671: gap of unknown length
* 54672 57546: contig of 2875 bp in length
* 57547 57646: gap of unknown length

* 57647 60740: contig of 3094 bp in length
* 60741 60840: gap of unknown length
* 60841 63916: contig of 3076 bp in length
* 63917 64016: gap of unknown length
* 64017 67000: contig of 2984 bp in length
* 67001 67100: gap of unknown length
* 67101 70219: contig of 3119 bp in length
* 70220 70319: gap of unknown length
* 70320 72950: contig of 2631 bp in length
* 72951 73050: gap of unknown length
* 73051 76802: contig of 3752 bp in length
* 76803 76902: gap of unknown length
* 76903 79990: contig of 3088 bp in length
* 79991 80090: gap of unknown length
* 80091 82745: contig of 2655 bp in length
* 82746 82845: gap of unknown length
* 82846 87725: contig of 4880 bp in length
* 87726 87825: gap of unknown length
* 87826 91178: contig of 3353 bp in length
* 91179 91278: gap of unknown length
* 91279 94533: contig of 3255 bp in length
* 94534 94633: gap of unknown length
* 94634 99871: contig of 5238 bp in length
* 99872 99971: gap of unknown length
* 99972 103742: contig of 3771 bp in length
* 103743 103842: gap of unknown length
* 103843 108877: contig of 5035 bp in length
* 108878 108977: gap of unknown length
* 108978 114019: contig of 5042 bp in length
* 114020 114119: gap of unknown length
* 114120 119231: contig of 5112 bp in length
* 119232 119331: gap of unknown length
* 119332 125649: contig of 6318 bp in length
* 125650 125749: gap of unknown length
* 125750 132687: contig of 6938 bp in length
* 132688 132787: gap of unknown length
* 132788 139741: contig of 6954 bp in length
* 139742 139841: gap of unknown length
* 139842 146349: contig of 6508 bp in length
* 146350 146449: gap of unknown length
* 146450 155325: contig of 8876 bp in length
* 155326 155425: gap of unknown length

Query Match 19.9%; Score 34.8; DB 2; Length 233269;
Best Local Similarity 62.8%; Pred. No. 1.1;
Matches 54; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy	39	ATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTGTGGCAGGCCGCTTCT	98
Db	95327	ATCTCCAAAGGTAACGTTATTAAAGGAAAGCTCCAAACAGCTGTGCCGTCACTCT	95386
Qy	99	CTTCAGCATGAGAAAGACAAAGGACCA 124	
Db	95387	TCTCAACAAGATAAAGACAGGTGCCA 95412	

RESULT 15
AC068103/c
LOCUS AC068103 144233 bp DNA linear HTG 14-JUN-2000
DEFINITION Homo sapiens chromosome 4 clone RP11-181I4 map 4, WORKING DRAFT
SEQUENCE, 38 unordered pieces.
ACCESSION AC068103
VERSION AC068103.2 GI:8516095
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 144233)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 4, clone RP11-181I4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 144233)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArelhano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczkzy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE

JOURNAL

COMMENT

Submitted (28-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 14, 2000 this sequence version replaced gi:7658404.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7833

Center clone name: 181_I_4

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 125189 bases at least Q40

Consensus quality: 133535 bases at least Q30

Consensus quality: 137574 bases at least Q20

Insert size: 160000; agarose-fp

Insert size: 140533; sum-of-contigs

Quality coverage: 3.1 in Q20 bases; agarose-fp

Quality coverage: 3.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 38 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1144: contig of 1144 bp in length
1145 1244: gap of 100 bp
1245 2299: contig of 1055 bp in length
2300 2399: gap of 100 bp
2400 3822: contig of 1423 bp in length
3823 3922: gap of 100 bp
3923 5190: contig of 1268 bp in length
5191 5290: gap of 100 bp
5291 5723: contig of 433 bp in length
5724 5823: gap of 100 bp
5824 7679: contig of 1856 bp in length
7680 7779: gap of 100 bp
7780 9186: contig of 1407 bp in length
9187 9286: gap of 100 bp
9287 11735: contig of 2449 bp in length
11736 11835: gap of 100 bp
11836 13249: contig of 1414 bp in length

* 13250 13349: gap of 100 bp
* 13350 14701: contig of 1352 bp in length
* 14702 14801: gap of 100 bp
* 14802 16706: contig of 1905 bp in length
* 16707 16806: gap of 100 bp
* 16807 18769: contig of 1963 bp in length
* 18770 18869: gap of 100 bp
* 18870 20539: contig of 1670 bp in length
* 20540 20639: gap of 100 bp
* 20640 23269: contig of 2630 bp in length
* 23270 23369: gap of 100 bp
* 23370 25610: contig of 2241 bp in length
* 25611 25710: gap of 100 bp
* 25711 27748: contig of 2038 bp in length
* 27749 27848: gap of 100 bp
* 27849 29177: contig of 1329 bp in length
* 29178 29277: gap of 100 bp
* 29278 30870: contig of 1593 bp in length
* 30871 30970: gap of 100 bp
* 30971 34125: contig of 3155 bp in length
* 34126 34225: gap of 100 bp
* 34226 36923: contig of 2698 bp in length
* 36924 37023: gap of 100 bp
* 37024 39589: contig of 2566 bp in length
* 39590 39689: gap of 100 bp
* 39690 42352: contig of 2663 bp in length
* 42353 42452: gap of 100 bp
* 42453 45755: contig of 3303 bp in length
* 45756 45855: gap of 100 bp
* 45856 49062: contig of 3207 bp in length
* 49063 49162: gap of 100 bp
* 49163 51707: contig of 2545 bp in length
* 51708 51807: gap of 100 bp
* 51808 56155: contig of 4348 bp in length
* 56156 56255: gap of 100 bp
* 56256 61528: contig of 5273 bp in length
* 61529 61628: gap of 100 bp
* 61629 66907: contig of 5279 bp in length
* 66908 67007: gap of 100 bp
* 67008 72627: contig of 5620 bp in length
* 72628 72727: gap of 100 bp
* 72728 77672: contig of 4945 bp in length
* 77673 77772: gap of 100 bp
* 77773 85211: contig of 7439 bp in length
* 85212 85311: gap of 100 bp
* 85312 92800: contig of 7489 bp in length
* 92801 92900: gap of 100 bp
* 92901 99694: contig of 6794 bp in length
* 99695 99794: gap of 100 bp
* 99795 108698: contig of 8904 bp in length
* 108699 108798: gap of 100 bp
* 108799 116661: contig of 7863 bp in length
* 116662 116761: gap of 100 bp
* 116762 128568: contig of 11807 bp in length
* 128569 128668: gap of 100 bp
* 128669 136426: contig of 7758 bp in length
* 136427 136526: gap of 100 bp
* 136527 144233: contig of 7707 bp in length.

FEATURES

source

Location/Qualifiers

1. .144233
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-181I4"
/clone_lib="RPCI-11 Human Male BAC"
1. .1144
/note="assembly_fragment"
1245. .2299
/note="assembly_fragment"
2400. .3822
/note="assembly_fragment"
3923. .5190

misc_feature

misc_feature

misc_feature

misc_feature

	Query Match	19.7%;	Score 34.4;	DB 2;	Length 144233;
	Best Local Similarity	54.8%;	Pred. No. 1.3;		
	Matches 68;	Conservative 0;	Mismatches 56;	Indels 0;	Gaps 0;
QY	9	AAC TTGAGCTT TGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGC	68		
Db	72574	ACCTGAAGCTCTGTGTTTTGAGGGATCAGACTTCCCTCACAGGTTACCAGAAAGGCCAGGT	72515		
QY	69	TCCACACCTCTGGCAGGGCCAGGGCCTTTCTCTTCAGCATGAGAAAGACAAGGGACACAGCAG	128		
Db	72514	AACACAAGCTCAATAAAGCTCAGGGAGCTTAGTGCTCACTCTGTGGAAACAAAATTGACCACTG	72455		
QY	129	AGTA	132		
Db	72454	AGAA	72451		

Search completed: July 29, 2003, 00:48:37
Job time : 491.369 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2003, 09:03:08 ; Search time 46.3409 Seconds
(without alignments)
8504.365 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708
Perfect score: 175
Sequence: 1 AGTCCAGGAACCTTGAGCTTT.....GCCTAGATAAACACCCAAA 175

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	175	100.0	1708	20	AAZ07516	Rat HICP polypepti
2	126.2	72.1	1734	20	AAX76488	Mouse WISP-2 prote
3	125.8	71.9	439	24	ABL59575	Rat OST23 gene fra
4	57	32.6	65	24	ABN30189	Rat spliced transc
5	30.8	17.6	463	22	AAK54634	Human haematologic
6	30.4	17.4	2155	23	ABL11587	Drosophila melanog
7	30.4	17.4	14091	23	ABL11586	Drosophila melanog
8	30	17.1	330	22	AAI91961	Human polynucleoti
9	30	17.1	422	22	AAF64359	Novel human polyn

10	30	17.1	483	22	ABA42323	Human breast cell
11	30	17.1	483	22	ABA52750	Human foetal liver
12	30	17.1	483	22	ABA22534	Probe #1000 for ge
13	30	17.1	483	22	AAK01005	Human brain expres
14	30	17.1	483	22	AAK26462	Human bone marrow
15	30	17.1	483	22	AAI11091	Probe #1024 for ge
16	30	17.1	483	22	AAI32356	Probe #1042 used t
17	30	17.1	483	22	AAI01010	Probe #1001 used t
18	30	17.1	483	24	ABS01052	Human genome-deriv
19	30	17.1	1096	22	ABA47464	Human breast cell
20	30	17.1	1096	22	ABA65355	Human foetal liver
21	30	17.1	1096	22	ABA32452	Probe #10918 for g
22	30	17.1	1096	22	AAK13765	Human brain expres
23	30	17.1	1096	22	AAK39508	Human bone marrow
24	30	17.1	1096	22	AAI20317	Probe #10250 for g
25	30	17.1	1096	22	AAI45522	Probe #14208 used
26	30	17.1	1096	22	AAI06020	Probe #6011 used t
27	30	17.1	1096	24	ABS13598	Human genome-deriv
28	30	17.1	1662	21	AAF22351	Human secreted pro
29	30	17.1	2016	22	ABA08988	Human glucocortico
30	29.8	17.0	2292	21	AAC74236	Human secreted pro
31	29.8	17.0	2304	21	AAC74249	Human secreted pro
32	29.8	17.0	1230025	20	AAX91990	Nucleotide sequenc
33	29.4	16.8	6236	24	AAD31108	Human transcriptio
34	29.4	16.8	25000	24	ABL68348	Kidney cancer rela
35	29.2	16.7	1721	22	AAF86139	Murine CDNA encodi
36	29.2	16.7	1984	19	AAV01596	Tobacco laccase cl
37	29	16.6	1568	24	ABK28657	Human CDNA encodin
38	29	16.6	1718	21	AAC76916	Human ORFX ORF2471
39	29	16.6	2658	23	AAS71055	DNA encoding novel
40	28.8	16.5	274	24	ABN76962	Human ORF1909 cDNA
41	28.8	16.5	5726	23	ABL02959	Drosophila melanog
42	28.8	16.5	8343	23	ABL02958	Drosophila melanog
43	28.6	16.3	3054	22	AAH98316	Human EST-derived
44	28.6	16.3	4758	23	AAS72577	DNA encoding novel
45	28.6	16.3	5955	23	AAS65525	DNA encoding novel

ALIGNMENTS

RESULT 1

AAZ07516

ID AAZ07516 standard; cDNA; 1708 BP.

XX

AC AAZ07516;

XX

DT 26-NOV-1999 (first entry)

XX

DE Rat HICP polypeptide encoding cDNA.

XX

KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;

KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.

XX

OS Rattus sp.

XX

PN WO9947556-A2.

XX

PD 23-SEP-1999.

XX

PF 18-MAR-1999; 99WO-US05999.

XX

PR 19-MAR-1998; 98US-0044273.

XX

(TUFT) TUFTS COLLEGE.

XX

PI Castellot JJ;

XX

DR WPI; 1999-562060/47.

DR

P-PSDB; AAY27434.

XX

PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein, used in methods to identify modulators or in diagnostic applications

PT

XX PS Claim 2; Fig 1; 108pp; English.

CC This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein.

CC Agents that stimulate or inhibit HICP protein activity or expression,

CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to

CC modulate cell-associated activity. HICP modulators can be used to treat

CC disorders characterized by aberrant HICP protein activity or expression.

CC Probes capable of hybridizing to HICP mRNA or antibodies specific for

CC HICP can be used to detect HICP activity in a biological sample. HICP

CC can be used to treat disorders, such as a cardiovascular or fibrotic

CC disorder, characterized by aberrant cell proliferation.

XX SQ Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Query Match 100.0%; Score 175; DB 20; Length 1708;

Best Local Similarity 100.0%; Pred. No. 1.4e-50;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

1534 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 1593

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAG 120

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

1594 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAG 1653

QY 121 GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

1654 GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 1708

RESULT 2

AAX76488

ID AAX76488 standard; DNA; 1734 BP.

XX AC AAX76488;

XX DT 06-AUG-1999 (first entry)

XX DE Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;

KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;

KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;

KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;

KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;

KW connective tissue disorder; catabolic state; inflammation;

KW testicular-related disorder; angiogenesis; immunological disorder; ss.

XX Mus sp.

XX WO9921998-A1.

XX 06-MAY-1999.

XX 29-OCT-1998; 98WO-US22991.

XX 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

XX (GETH) GENENTECH INC.

XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

XX WPI; 1999-337420/28.

DR P-PSDB; AAY17651.

XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

PT Example 2; Page 178-179; 284pp; English.

XX CC The present invention describes Wnt-1 induced secreted polypeptides,

CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2

CC and WISP-3 have homology to connective tissue growth factor (CTGF).

CC Products from the present invention can be used to treat WISP-related

CC disorders such as breast, ovarian, and colon cancer or melanoma. The

CC products can be used to treat arteriosclerosis. The products can also be

CC used to treat other diseases e.g. benign and malignant tumours,

CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,

CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and

CC blastocoeleic disorders, haematopoiesis-related disorders, tissue-growth

CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney

CC disorders, bone-related disorders such as osteoporosis, trauma such as

CC burns, incisions, and other wounds, connective tissue disorders,

CC catabolic states, testicular-related disorders, and inflammatory,

CC angiogenic and immunologic disorders including arteriosclerosis. The

CC products can also be used for detection and diagnosis especially of

CC individuals with neoplastic cell growth or proliferation. The products

CC can be used in the production of transgenic or knock-out animals.

CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing

CC cells.

XX SQ Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;

Query Match 72.1%; Score 126.2; DB 20; Length 1734;

Best Local Similarity 86.3%; Pred. No. 1.4e-33;

Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

1536 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 1595

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAG 120

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

1596 AAG-AGGCTCCACACCTCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAGACAAG 1654

QY 121 GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

1655 AACAGTAGAGTACCTCTCTGGAGGACTGGCCCGGTCTGGAATAAACACCCAAA 1709

RESULT 3

ABL59575

ID ABL59575 standard; DNA; 439 BP.

XX AC ABL59575;

XX DT 17-JUL-2002 (first entry)

XX DE Rat OST23 gene fragment SEQ ID NO:23.

XX Rat; OST; osteoregenerative; parathyroid hormone; tibiae; osteopathic;

KW gene therapy; parathyroid hormone receptor ligand; bone disorder;

KW bone formation disorder; bone resorption disorder; osteoporosis;

KW osteopenia; osteopetrosis; gene; ds.

XX Rattus sp.

XX WO200224943-A2.

XX 28-MAR-2002.

XX 19-SEP-2001; 2001WO-US29548.

XX 19-SEP-2000; 2000US-233579P.

XX (CURA-) CURAGEN CORP.

PA (GLAX) GLAXO GROUP LTD.

XX Horesovsky GJ, Noel ES, Raha D;

PI WPI; 2002-401989/43.

XX

PT Identifying parathyroid hormone receptor ligands and osteoregenerative
PT agents involves detecting the expression of nucleic acids which are
PT regulated by parathyroid hormone
XX
PS Claim 45; Page 22; 90pp; English.
XX
CC The present invention describes a method (M1) for identifying parathyroid
CC hormone receptor ligands (I) and osteoregenerative agents by contacting a
CC test cell population (CP) comprising cells expressing nucleic acid
CC sequences (S) of OST1-47 and 48, with a test agent, measuring nucleic
CC acid sequence expression, comparing it with reference CP and identifying
CC the ligand and agent by the difference in expression levels. Also
CC described is a method (M2) for treating a bone disorder in a subject by
CC administering to the subject an agent that modulates the expression or
CC activity of (S). (I) have osteopathic activities, and can be used to
CC modulate the expression of OST 1-48, and can also be used in gene
CC therapy. (M1) is useful for identifying parathyroid hormone receptor
CC ligands (I) or osteoregenerative agents. OST gene sequence can be used
CC for assessing the osteoregenerative activity of a test agent in a
CC subject, and for diagnosing or determining the susceptibility to bone
CC disorder and assessing efficacy of a treatment of a bone disorder in a
CC subject, human or rodent. (M2) is useful for treating a bone disorder
CC including bone formation disorder or bone resorption disorder (e.g.
CC osteoporosis, osteopenia and osteopetrosis). OST polypeptides are useful
CC as immunogens to raise anti-OST antibodies (II). Parathyroid hormone
CC receptor ligands (I) identified by (M1) are useful for treating bone
CC disorders. The present sequence represents an OST23 gene fragment
CC isolated from rat tibiae, from the present invention.
XX
SQ Sequence 439 BP; 128 A; 108 C; 99 G; 104 T; 0 other;

Query Match 71.9%; Score 125.8; DB 24; Length 439;
Best Local Similarity 96.8%; Pred. No. 1.1e-33;
Matches 150; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 AGTCCAGGAACTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 285 AGTCCAGGAACTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 344
QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCCCTTTCTCTTTCAGCATGAGAAACAAGG 120
|||||
Db 345 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCCCTTTCTCTTTCAGCATGAGAAACAAGG 404
QY 121 G--ACAGCAGAGTAC-TCTCCTCTGGAGGACTAGT 152
|
Db 405 GGACCAGCAGAGTACTTTTCCTCTGGAGGACTAGT 439

RESULT 4
ABN30189
ID ABN30189 standard; DNA; 65 BP.
XX ABN30189;
XX
DT 15-JUL-2002 (first entry)
XX
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:2937.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Rattus norvegicus.
XX
XX WO200210449-A2.
PN
XX
PD 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-IB01903.
PF
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.

XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
PS Example 1; SEQ ID 2937; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 21 A; 20 C; 13 G; 11 T; 0 other;

Query Match 32.6%; Score 57; DB 24; Length 65;
Best Local Similarity 92.3%; Pred. No. 4.4e-10;
Matches 60; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 26 TTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTGGCAGG 85
|||||
Db 1 TTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTTAACAGC 60
QY 86 CCAGG 90
|
Db 61 CAGGG 65

RESULT 5
AAK54634
ID AAK54634 standard; cDNA; 463 BP.
XX
AC AAK54634;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human haematological malignancy-related antigen coding sequence #359.
XX
KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
OS Homo sapiens.
XX
PN WO200164886-A2.
XX
PD 07-SEP-2001.
XX
PF 01-MAR-2001; 2001WO-US07272.

XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX P-PSDB; ABB67483.
DR
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 29240; 2lpp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 14091 BP; 3932 A; 3192 C; 3264 G; 3703 T; 0 other;
Query Match 17.4%; Score 30.4; DB 23; Length 14091;
Best Local Similarity 59.1%; Pred. No. 7.6;
Matches 52; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 84 GGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAAGGACAGCAGAGTACTCTCCTCTGG 143
Db 8287 GACCACGAACACGCTCTCCGGCTCGATGAGGAATATGGGCAGCGGTATTACTCGCCTCTCC 8228
QY 144 AGGACTAGTCTAGCCTAGATAAACACC 171
Db 8227 GGGACTCAGTTTGCCGCTGTTAAACTCC 8200
RESULT 8
AAI91961/c
ID AAI91961 standard; cDNA; 330 BP.
XX
AC AAI91961;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 12021.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO12030.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 1; SEQ ID NO 12021; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 330 BP; 106 A; 91 C; 70 G; 63 T; 0 other;
Query Match 17.1%; Score 30; DB 22; Length 330;
Best Local Similarity 57.4%; Pred. No. 2.2;
Matches 54; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 9 AACTTGAGCTTGTATTTCAGGAATGCACATCTCTTAAGCACCTCGCAAAACAGGAAGGC 68
Db 248 AATTTTCATCTGTGTCCATTTCAGTAGTCTTTCAGCAGTTGGGCAGCCCCCAGATGAGGATGC 189
QY 69 TCCACACCTCTGGCAGGCCAGGCCCTTTCTCTTC 102
Db 188 TGGCTGCGATGGGGAGCACACAGGGTCTTTCAGCTC 155
RESULT 9
AAF64359/c
ID AAF64359 standard; cDNA; 422 BP.
XX
AC AAF64359;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 115.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:55:06 ; Search time 9.52094 Seconds
(without alignments)
5636.882 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708
Perfect score: 175
Sequence: 1 AGTCCAGGAACCTTGAGCTTT.....GCTAGATAAACACCCCAA 175

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	126.2	72.1	1734	4 US-09-182-145-17	Sequence 17, Appl
2	126.2	72.1	1734	4 US-09-182-145-18	Sequence 18, Appl
3	27.8	15.9	10825	3 US-08-652-265-1	Sequence 1, Appli
4	27.8	15.9	10825	3 US-08-652-265-3	Sequence 3, Appli
5	27.8	15.9	10825	3 US-08-652-265-5	Sequence 5, Appli
6	27.8	15.9	10825	3 US-08-652-265-7	Sequence 7, Appli
7	27.8	15.9	10825	3 US-08-834-497A-1	Sequence 1, Appli
8	27.8	15.9	10825	3 US-08-834-497A-3	Sequence 3, Appli
9	27.8	15.9	10825	3 US-08-834-497A-5	Sequence 5, Appli
10	27.8	15.9	10825	3 US-09-503-444A-1	Sequence 7, Appli
11	27.8	15.9	10825	4 US-09-503-444A-3	Sequence 1, Appli
12	27.8	15.9	10825	4 US-09-503-444A-5	Sequence 3, Appli
13	27.8	15.9	10825	4 US-09-503-444A-7	Sequence 5, Appli
14	27.8	15.9	10825	4 US-09-503-444A-9	Sequence 7, Appli
15	27.8	15.9	12146	4 US-09-277-457-27	Sequence 27, Appl
16	27.8	15.9	246240	2 US-08-724-394A-20	Sequence 20, Appl
17	27.8	15.9	246240	2 US-08-724-394A-21	Sequence 22, Appl
18	27.8	15.9	246240	2 US-08-724-394A-22	Sequence 21, Appl
19	27.8	15.4	13011	2 US-08-791-849A-14	Sequence 14, Appl
20	26.8	15.3	11236	1 US-07-853-913-1	Sequence 1, Appli
21	26.6	15.2	519	1 US-08-438-753B-43	Sequence 43, Appl
22	26.6	15.2	519	1 US-08-443-883A-43	Sequence 43, Appl
23	26.6	15.2	519	2 US-08-631-328-43	Sequence 43, Appl
24	26.6	15.2	519	2 US-08-455-524B-43	Sequence 43, Appl
25	26.6	15.2	519	2 US-08-455-021B-43	Sequence 43, Appl
26	26.6	15.2	519	4 US-09-045-467-43	Sequence 43, Appl
27	26.6	15.2	588	1 US-08-438-753B-11	Sequence 11, Appl

28	26.6	15.2	588	1 US-08-443-883A-11	Sequence 11, Appl
29	26.6	15.2	588	2 US-08-631-328-11	Sequence 11, Appl
30	26.6	15.2	588	2 US-08-455-524B-11	Sequence 11, Appl
31	26.6	15.2	588	2 US-08-455-021B-11	Sequence 11, Appl
32	26.6	15.2	588	4 US-09-045-467-11	Sequence 11, Appl
33	26.6	15.2	1023	3 US-09-188-930-22	Sequence 22, Appl
34	26.6	15.2	1697	3 US-09-188-930-264	Sequence 264, App
35	26.2	15.0	1023	1 US-08-698-551-7	Sequence 7, Appli
36	26.2	15.0	1023	2 US-08-602-228-7	Sequence 7, Appli
37	26.2	15.0	1023	2 US-08-649-341A-7	Sequence 7, Appli
38	26.2	15.0	1023	2 US-08-494-440B-7	Sequence 7, Appli
39	26.2	15.0	1023	2 US-08-533-901B-7	Sequence 7, Appli
40	26.2	15.0	1023	2 US-08-839-032A-7	Sequence 7, Appli
41	26.2	15.0	1023	2 US-08-839-031A-7	Sequence 7, Appli
42	26.2	15.0	1023	4 US-09-185-258C-7	Sequence 7, Appli
43	26.2	15.0	1023	5 PCT-US92-01196-5	Sequence 5, Appli
44	26.2	15.0	1023	5 PCT-US95-12724-7	Sequence 7, Appli
45	26.2	15.0	1650	4 US-08-456-640-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-182-145-17
; Sequence 17, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-182-145-17

Query Match	72.1%	Score	126.2;	DB	4;	Length	1734;
Best Local Similarity	86.3%;	Pred. No.	1.6e-32;				
Matches	151;	Conservative	0;	Mismatches	23;	Indels	1;
Gaps	1;						
QY	1	AGTCCAGGAACCTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTTAAGCACTCCCAAAAC	60				
Db	1536	AGTCCAGGAACCTTGAGCTTTGTATTTGTAAAAATACACATCTCTTTAAATGCTCACAAAGC	1595				
QY	61	AGGAAGGCTCCACACTCTGGCAGGCCAGGGCCCTTCTTTCAGCATGAGAAAGACAAGG	120				
Db	1596	AAG-AGGCTCCACACTCTGGCAGGCCAGGGCCCTTCTTTCAGCATGAGAGACAAGG	1654				
QY	121	GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGACCTAGATAAACACCCAAA	175				
Db	1655	AACAGTAGAGTACCCTCTCTGGAGGACTGGCCCGTCTGGAATAAACACCCAAA	1709				

RESULT 2

Db 2609 AGGCCAAGGAGAGCAGATTCTCTGAGCTCAGGAGTTCAAGACCAGCTGGGCAACACAGCA 2668
Qy 173 AAA 175
Db 2669 AAA 2671

RESULT 4

US-08-652-265-3
; Sequence 3, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:

NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
OTHER INFORMATION: mutation"
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
OTHER INFORMATION: gene 24dl allele"
FEATURE:
NAME/KEY: -
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: 24dl allele cDNA (SEQ ID NO:10)"
FEATURE:
NAME/KEY: -
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d2(C) allele (SEQ ID NO:41)"
FEATURE:

NAME/KEY: -
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24dl(A) allele (SEQ ID NO:21)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(5834, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24dl
US-08-652-265-3
Query Match 15.9%; Score 27.8; DB 3; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 113 AGACAAGGGACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCC 172
Db 2609 AGGCCAAGGAGAGCAGATTCTCTGAGCTCAGGAGTTCAAGACCAGCTGGGCAACACAGCA 2668
Qy 173 AAA 175
Db 2669 AAA 2671

RESULT 5

US-08-652-265-5
; Sequence 5, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,


```

; Sequence 1, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/632,673
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5898)
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: /note= "No. 6140305mal or wild-type (un
; OTHER INFORMATION: Hereditary Hemochromatosis (HH) gene
; OTHER INFORMATION: allele"
; FEATURE:
; NAME/KEY: -
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) allele
; OTHER INFORMATION: cDNA (SEQ ID NO:9)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3852..3891

```

```

; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d2(C)
; OTHER INFORMATION: allele (SEQ ID NO:41)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d1(G)
; OTHER INFORMATION: allele (SEQ ID NO:20)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3872, "c")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d2
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3878, "a")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d7
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "g")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d1
;
US-08-834-497A-1

Query Match 15.9%; Score 27.8; DB 3; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0;

QY 113 AGACAAGGGACAGACAGAGTAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGAATA
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2609 AGCCCAAGGAGACAGAGATTCTCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAA

QY 173 AAA 175
|||
Db 2669 AAA 2671

RESULT 8
US-08-834-497A-3
; Sequence 3, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSEO for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997

```

```
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/652,265
;; FILING DATE: 23-MAY-1996
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/632,673
;; FILING DATE: 16-APR-1996
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/630,912
;; FILING DATE: 04-APR-1996
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 8907-0056-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-493-4935
;; TELEFAX: 650-493-5556
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10825 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
;; LOCATION: 6040..6153, 7107..7147)
;; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
;; OTHER INFORMATION: mutation"
;; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
;; OTHER INFORMATION: gene 24d1 allele"
;; FEATURE:
;; NAME/KEY: -
;; LOCATION: 140..7319
;; OTHER INFORMATION: /note= "start and stop positions for
;; OTHER INFORMATION: 24d1 allele cDNA (SEQ ID NO:10)"
;; FEATURE:
;; NAME/KEY: -
;; LOCATION: 3852..3891
;; OTHER INFORMATION: /note= "start and stop positions for
;; OTHER INFORMATION: genomic sequence surrounding variant
;; OTHER INFORMATION: for 24d2(C) allele (SEQ ID NO:41)"
;; FEATURE:
;; NAME/KEY: -
;; LOCATION: 5507..6023
;; OTHER INFORMATION: /note= "start and stop positions for
;; OTHER INFORMATION: genomic sequence surrounding variant
;; OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
;; FEATURE:
;; NAME/KEY: allele
;; LOCATION: replace(5834, "a")
;; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
;; OTHER INFORMATION:
;; OTHER INFORMATION: /label= 24d1
US-08-834-497A-3
```

```
Query Match 15.9%; Score 27.8; DB 3; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 113 AGACAAGGGACAGCAGAGTAGTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCC 172
||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 2609 AGCCCAAGGAGAGCAGGATTCCTGAGCTCAGGAGTCAAGACCAGCCTGGCAACACAGCA 2668

QY 173 AAA 175
|||
Db 2669 AAA 2671
```

```
RESULT 9
US-08-834-497A-5
; Sequence 5, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/632,673
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
; OTHER INFORMATION: mutation"
; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
; OTHER INFORMATION: gene 24d2 allele"
; FEATURE:
; NAME/KEY: -
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: 24d2 allele cDNA (SEQ ID NO:11)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d2(C) allele (SEQ ID NO:41)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
US-08-834-497A-3
```



```

; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d2(C)
; OTHER INFORMATION: allele (SEQ ID NO:41)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d1(G)
; OTHER INFORMATION: allele (SEQ ID NO:20)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3872, "c")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d2
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3878, "a")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d7
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "g")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d1
;
US-09-503-444A-1

Query Match 15.9%; Score 27.8; DB 4; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0;

QY 113 AGACAAGGGACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCTAGAATA
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2609 AGCCCAAGAGAGAGAGATTCTCTGAGCTCAGGAGTTCAAGACCAGCTGGGCAA

QY 173 AAA 175
|||
Db 2669 AAA 2671

RESULT 12
US-09-503-444A-3
; Sequence 3, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect Version 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503.444A
; FILING DATE: 14-Feb-2000
; CLASSIFICATION:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/652,265
; FILING DATE: 23-May-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/632,673
; FILING DATE: 16-Apr-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,912
; FILING DATE: 04-Apr-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0088-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
; OTHER INFORMATION: mutation"
; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
; OTHER INFORMATION: gene 24d1 allele"
; FEATURE:
; NAME/KEY: -
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: 24d1 allele cDNA (SEQ ID NO:10)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d2(C) allele (SEQ ID NO:41)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION: /label= 24d1
; US-09-503-444A-3
;
Query Match 15.9%; Score 27.8; DB 4; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
;
QY 113 AGACAAGGGACAGCAGAGTAGTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCC 172
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2609 AGGCCAAGGAGAGCAGAGATTCTGTAGCTCAGGAGTCAAGACCAGCCTGGGCAACACAGCA 2668
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 173 AAA 175
| | |
Db 2669 AAA 2671
| | |
;
RESULT 13
US-09-503-444A-5
```

```

; Sequence 5, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect Version 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,444A
; FILING DATE: 14-Feb-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/652,265
; FILING DATE: 23-May-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/632,673
; FILING DATE: 16-Apr-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,912
; FILING DATE: 04-Apr-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0088-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
; OTHER INFORMATION: mutation"
; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
; OTHER INFORMATION: gene 24d2 allele"
; FEATURE:
; NAME/KEY: -
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: 24d2 allele cDNA (SEQ ID NO:11)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
; FEATURE:
```

NAME/KEY: 5507..6023
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d1(G) allele (SEQ ID NO:20)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(3872, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
OTHER INFORMATION: /label= 24d2
US-09-503-444A-5

Query Match 15.9%; Score 27.8; DB 4; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 113 AGACAAGGGACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCC 172
Db 2609 AGGCCAAGGAGAGCAGAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACACAGCA 2668
QY 173 AAA 175
Db 2669 AAA 2671

RESULT 14
US-09-503-444A-7
Sequence 7, Application US/09503444A
Patent No. 6228594
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Wordperfect Version 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/503,444A
FILING DATE: 14-Feb-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/652,265
FILING DATE: 23-May-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,673
FILING DATE: 16-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
FILING DATE: 04-Apr-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis
OTHER INFORMATION: and 24d2 mutations"
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
OTHER INFORMATION: gene containing a combination of both
OTHER INFORMATION: 24d1 and 24d2 alleles"
FEATURE:
NAME/KEY: -
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: CDNA containing a combination of both
OTHER INFORMATION: 24d1 and 24d2 alleles
OTHER INFORMATION: (SEQ ID NO:12)"
FEATURE:
NAME/KEY: -
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
FEATURE:
NAME/KEY: -
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(3872, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(5834, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
OTHER INFORMATION: /label= 24d1
US-09-503-444A-7
Query Match 15.9%; Score 27.8; DB 4; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 113 AGACAAGGGACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCC 172
Db 2609 AGGCCAAGGAGAGCAGAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACACAGCA 2668
QY 173 AAA 175
Db 2669 AAA 2671
RESULT 15
US-09-277-457-27
Sequence 27, Application US/09277457
Patent No. 6355425
GENERAL INFORMATION:
APPLICANT: Rothenberg, Barry E.
APPLICANT: Sawada-Hirai, Ritsuko
APPLICANT: Barton, James G.
TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
FILE REFERENCE: 10653/002001


```

; CURRENT APPLICATION NUMBER: US/09/277,457
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12146
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-277-457-27

```

Query Match 15.9%; Score 27.8; DB 4; Length 12146;
Best Local Similarity 65.1%; Pred. No. 12;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY	113	AGACAAGGGACAGCAGAGTACTCTCCCTCTGGAGGACTAGTCTAGCCTAGAATAAACACCC	172
Db	3496	AGGCCAAGGAGACAGATTCCTTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACACAGCA	3555
QY	173	AAA	175
Db	3556	AAA	3558

Search completed: July 28, 2003, 21:35:59
Job time : 11.7209 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:36:41 ; Search time 43.1391 Seconds
(without alignments)
8368.853 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708
Perfect score: 175
Sequence: 1 AGTCCAGGAAGCTTGAGCTTT.....GCCTAGATAAACACCCAAA 175

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1.*
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3.*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	175	100.0	1708	15	US-10-010-408-1
2	126.2	72.1	1734	15	US-10-112-267-17
3	126.2	72.1	1734	15	US-10-112-267-18
4	125.8	71.9	439	12	US-09-956-622A-23
5	31.2	17.8	707	15	US-10-027-632-12045
6	30.8	17.6	463	11	US-09-796-692-359
7	30.8	17.6	463	11	US-09-796-692-4928
8	30.8	17.6	463	15	US-10-040-862-359
9	30.8	17.6	463	15	US-10-040-862-4928
10	30	17.1	483	10	US-09-864-761-1000
11	30	17.1	1096	10	US-09-864-761-1772
12	29.4	16.8	1753	15	US-10-161-803-23
13	29	16.6	624	15	US-10-027-632-205550
14	29	16.6	851	15	US-10-027-632-110489
15	28.8	16.5	431	11	US-09-560-863-222
16	28.8	16.5	659158	11	US-09-771-208-20

17	28.6	16.3	570	15	US-10-027-632-281405	Sequence 281405,
18	28.6	16.3	627	15	US-10-027-632-189371	Sequence 189371,
19	28.6	16.3	627	15	US-10-027-632-189372	Sequence 189372,
20	28.4	16.2	423	11	US-09-960-352-3889	Sequence 3889, Ap
21	28.4	16.2	429	10	US-09-897-214-6	Sequence 6, Appli
22	28.4	16.2	430	11	US-09-960-352-12168	Sequence 12168, A
23	28.4	16.2	444	15	US-10-106-698-1852	Sequence 1852, Ap
24	28.4	16.2	1653	10	US-09-897-214-9	Sequence 9, Appli
25	28.2	16.1	3186778	15	US-10-027-632-174961	Sequence 174961,
26	28	16.0	419	11	US-09-960-352-12167	Sequence 12167, A
27	28	16.0	439	11	US-09-983-965-3909	Sequence 3909, Ap
28	28	16.0	698	15	US-10-027-632-21981	Sequence 21981, A
29	28	16.0	698	15	US-10-027-632-21982	Sequence 21982, A
30	28	16.0	2432	15	US-10-027-632-101680	Sequence 101680,
31	28	16.0	25701	12	US-09-764-891-9766	Sequence 9766, Ap
32	28	16.0	25758	12	US-09-764-891-9765	Sequence 9765, Ap
33	27.8	15.9	611	15	US-10-027-632-296356	Sequence 296356,
34	27.8	15.9	624	15	US-10-027-632-5700	Sequence 5700, Ap
35	27.8	15.9	951	11	US-09-886-055-134	Sequence 134, App
36	27.8	15.9	951	12	US-09-804-291-134	Sequence 134, App
37	27.8	15.9	5982	15	US-10-016-634A-25	Sequence 25, Appl
38	27.8	15.9	12146	9	US-09-981-606-27	Sequence 27, Appl
39	27.8	15.9	40433	11	US-09-880-107-3327	Sequence 3327, Ap
40	27.8	15.9	235033	15	US-10-301-844-1	Sequence 1, Appli
41	27.8	15.9	237326	15	US-10-301-844-2	Sequence 2, Appli
42	27.6	15.8	440	12	US-09-918-995-29616	Sequence 29616, A
43	27.6	15.8	779	11	US-09-731-872-131	Sequence 131, App
44	27.4	15.7	529	12	US-09-918-995-9209	Sequence 9209, Ap
45	27.4	15.7	566	15	US-10-027-632-284014	Sequence 284014,

ALIGNMENTS

RESULT 1

US-10-010-408-1
; Sequence 1, Application US/10010408
; Publication No. US20020165185A1
; GENERAL INFORMATION:
; APPLICANT: John J. Castellet, Jr.
; TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules and Uses Therefor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/010,408
; FILING DATE: 07-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/044,273
; FILING DATE: March 19, 1998
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MBI-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-956-622A-23

Query Match 71.9%; Score 125.8; DB 12; Length 439;
Best Local Similarity 96.8%; Pred. No. 2.6e-35;
Matches 150; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 AGTCCAGGAAGTGGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 285 AGTCCAGGAAGTGGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 344
|||||

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCCTTCTCTTCAGATGAGAAAGACAAGG 120
|||||
Db 345 AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCCTTCTCTTCAGATGAGAAAGACAAGG 404
|||||

QY 121 G--ACAGCAGAGTAC-TCTCCTCTGGAGGACTAGT 152
|
Db 405 GGACCAGCAGAGTACTTTTCTCTCTGGAGGACTAGT 439
|||||

RESULT 5
US-10-027-632-12045/C
; Sequence 12045, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12045
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-12045

Query Match 17.8%; Score 31.2; DB 15; Length 707;
Best Local Similarity 54.3%; Pred. No. 0.47;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 36 CACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTCTGGCAGGCCAGGCCTT 95
|||||
Db 686 CTCAGCTCCCAGAAAGTCTCAGAGGGTGAGGGTCCCACATCTCTCTGAGGACAGGCCCTA 627
|||||

QY 96 TCTCTTCAGCATGAGAAAGACRAGGGACAGAGTACTCTCTCTGGAGGACTAG 151
|||
Db 626 GCTAACGAGTACAGAAACCAGGGCCTAAAGTCCCAATCCCACAGAGGCTGG 571
|||||

RESULT 6
US-09-796-692-359
; Sequence 359, Application US/09796692
; Publication No. US20020198362A1

; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 359
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(463)
; OTHER INFORMATION: n = A,T,C or G
US-09-796-692-359

Query Match 17.6%; Score 30.8; DB 11; Length 463;
Best Local Similarity 53.3%; Pred. No. 0.57;
Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1 AGTCCAGGAAGTGGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 265 AGTCCTTCAACTTCTGTTCTTGTCTAAGAGGTGTGCAAAATGAATTACCAACTGGGTAACC 324
|||||

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCCTTCTCTTCAGATGAGAAAGACAAGG 120
|||||
Db 325 TTGAATGAGTTAATCAAATTCAGGCCATCGTCCAGCCACCAACAGTCAAAAGGCCAAGG 384
|||||

QY 121 GA 122
Db 385 AA 386
|

RESULT 7
US-09-796-692-4928
; Sequence 4928, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200

```

Db          385 AA 386:

RESULT 8
US-10-040-862-359
; Sequence 359, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0.0
; SEQ ID NO 359
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(463)
; OTHER INFORMATION: n = A,T,C or G
US-10-040-862-359

```

	Query Match	17.6%; Best Local Similarity Matches 65;	Score 30.8; Pred. No. 0.57; Conservative 0;	DB 15; Mismatches 57;	Length 463; Indels 0;	Gaps 0
QY	1	AGTCAGGAACCTTGAGCTTTGTATTTT	CAGGAATGCACATCTCTTAAGCACTCGCAAAC	60		
Db	265	AGTCCTTCAACTTCTGTTTTCTTGCTAACAGGTGTGCAATGAATTACCAACTGGGTAACC	324			
QY	61	AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCCCTTTCTCTT	CACCATGAGAAGACAAGG	120		
Db	325	TTGAATGAGTTAATCAAATCCAGGCCATCGTCCACGCCACCAACAGTAGTCAAGGCCAAGG	384			
QY	121	GA 122				
Db	385	AA 386				

RESULT 9

US-10-040-862-4928
; Sequence 4928, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4928
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (6)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (43)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (45)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (47)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (108)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (110)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure

; LOCATION: (420)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (448)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-4928

Query Match 17.6%; Score 30.8; DB 15; Length 463;
Best Local Similarity 53.3%; Pred. No. 0.57;
Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY 1 AGTCCAGGAAC TTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
||||| ||||| | | | |||| |||| | || | ||| | |||
Db 265 AGTCCTCAACTTCTGTTCTTGCTAACAGGTGTGCAAAATGAATTACCAACTGGGTAACC 324
QY 61 AGGAGGCTCCACACCTCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGAAAGACAAGG 120
||||| | | | ||||| | | | |||||
Db 325 TTGAATGAGTTAATCAAATTCAGGCCATCGTCCACGCCAACCAACAGTCAAGAGGCCAAGG 384
QY 121 GA 122
Db 385 AA 386

RESULT 10

US-09-864-761-1000
; Sequence 1000, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30

Fri Aug 1 10:44:12 2003

us-10-010-408-1_copy_1534_1708.rnppb

Page 8

Search completed: July 28, 2003, 21:31:58
Job time : 47.1391 secs

GenCore version 5.1.6
Copyright (c) 1993.- 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 28, 2003, 18:02:17 ; Search time 310.315 Seconds
(without alignments)
9133.322 Million cell updates/sec
Title: US-10-010-408-1_COPY_1534_1708
Perfect score: 175
Sequence: 1 AGTCCAGGAAGCTTGAGCTTT.....GCCTAGATAAACACCCAAA 175

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				EST: *	
1:	em_estba:	*		1:	em_estba:
2:	em_esthum:	*		2:	em_esthum:
3:	em_estin:	*		3:	em_estin:
4:	em_estnu:	*		4:	em_estnu:
5:	em_estov:	*		5:	em_estov:
6:	em_estpl:	*		6:	em_estpl:
7:	em_estro:	*		7:	em_estro:
8:	em_htc:	*		8:	em_htc:
9:	gb_est1:	*		9:	gb_est1:
10:	gb_est2:	*		10:	gb_est2:
11:	gb_htc:	*		11:	gb_htc:
12:	gb_est3:	*		12:	gb_est3:
13:	gb_est4:	*		13:	gb_est4:
14:	gb_est5:	*		14:	gb_est5:
15:	em_estfun:	*		15:	em_estfun:
16:	em_estom:	*		16:	em_estom:
17:	gb_gss:	*		17:	gb_gss:
18:	em_gss_hum:	*		18:	em_gss_hum:
19:	em_gss_inv:	*		19:	em_gss_inv:
20:	em_gss_pln:	*		20:	em_gss_pln:
21:	em_gss_vrt:	*		21:	em_gss_vrt:
22:	em_gss_fun:	*		22:	em_gss_fun:
23:	em_gss_mam:	*		23:	em_gss_mam:
24:	em_gss_mus:	*		24:	em_gss_mus:
25:	em_gss_other:	*		25:	em_gss_other:
26:	em_gss_pro:	*		26:	em_gss_pro:
27:	em_gss_rod:	*		27:	em_gss_rod:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	173.4	99.1	762	14 BQ195526	BQ195526 UI-R-CN1-
2	126.2	72.1	418	9 AI225477	AI225477 ue88b01.y
3	124.6	71.2	313	10 BB367824	BB367824 BB367824
4	123	70.3	315	10 BB374499	BB374499 BB374499
5	119.8	68.5	307	10 BB220676	BB220676 BB220676
6	119.8	68.5	337	10 BB319151	BB319151 BB319151

7	119.8	68.5	345	10	BB222795	BB222795
8	113.4	64.8	281	10	BB309266	BB309266
9	113.4	64.8	316	10	BB518921	BB518921
10	110.2	63.0	218	10	BB213539	BB213539
11	110.2	63.0	289	10	BB186801	BB186801
12	110.2	63.0	290	10	BB515213	BB515213
13	104.8	59.9	369	10	BB792544	BB792544
14	100.2	57.3	324	10	BB558051	BB558051
15	74.4	42.5	792	12	BF138093	BF138093
16	33.6	19.2	675	17	AZ574808	AZ574808
17	33.6	19.2	712	13	BI147316	BI147316
18	33.4	19.1	360	17	AQ345221	AQ345221
19	33.4	19.1	506	10	BB701454	BB701454
20	33.2	19.0	469	9	AL800293	AL800293
21	32.2	18.4	317	12	BF077992	BF077992
22	32.2	18.4	441	10	BB775647	BB775647
23	32.2	18.4	589	12	BF712425	BF712425
24	32.2	18.4	620	10	AV751782	AV751782
25	32	18.3	677	17	AZ574825	AZ574825
26	32	18.3	682	17	AZ574814	AZ574814
27	32	18.3	744	17	CNS04PR1	AL301654 Tetraodon
28	32	18.3	1089	17	CNS031YH	AL246194 Tetraodon
29	31.8	18.2	790	9	AL820075	AL820075
30	31.6	18.1	799	13	BI757228	BI757228
31	31.4	17.9	1135	14	BQ278875	BQ278875
32	31.2	17.8	412	10	AW694655	AW694655
33	31.2	17.8	454	9	AL379783	AL379783
34	31	17.7	263	9	AV063522	AV063522
35	31	17.7	327	12	BF549593	BF549593
36	31	17.7	373	17	AQ120776	AQ120776
37	31	17.7	434	13	BM256378	BM256378
38	31	17.7	506	10	BE485090	BE485090
39	31	17.7	639	13	BJ254458	BJ254458
40	31	17.7	691	17	AQ752684	AQ752684
41	31	17.7	697	17	AZ127785	AZ127785
42	31	17.7	740	14	BQ210400	BQ210400
43	31	17.7	799	17	AZ185214	AZ185214
44	31	17.7	929	14	BQ425393	BQ425393
45	30.8	17.6	418	14	W97990	W97990

ALIGNMENTS

RESULT 1	BQ195526/c	BQ195526	762 bp	mRNA	linear	EST 30-APR-2002
LOCUS	UI-R-CN1-cmq-k-07-0-UI.s1	UI-R-CN1	UI-R-CN1	Rattus norvegicus	cdna clone	
DEFINITION	UI-R-CN1-cmq-k-07-0-UI 3', mRNA sequence.					
ACCESSION	BQ195526					
VERSION	BQ195526.1	GI:20371077				
KEYWORDS	EST.					
SOURCE	Norway rat.					
ORGANISM	Rattus norvegicus					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
REFERENCE	1 (bases 1 to 762)					
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.					
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery					
JOURNAL	Genome Res. 6 (9), 791-806 (1996)					
MEDLINE	97044477					
COMMENT	Contact: Soares, MB Program for Rat Gene Discovery and Mapping University of Iowa 451 Eckstein Medical Research Building Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: msoares@blue.weeg.uiowa.edu The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A					

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dr track served to identify it as a clone from the normalized cervix library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) Seq primer: M13 Forward POLYA=Yes.

FEATURES source

Location/Qualifiers

1. .762
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CN1-cmq-k-07-0-UI"
/clone_lib="UI-R-CN1"
/dev_stage="adult"

/lab_host="DHL0B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CN1 library is a subtracted library derived from the following pool of seven normalized rat libraries: normalized rat seminal vesicles, normalized rat penis, normalized rat bladder, normalized rat cervix, normalized rat brown adipose, normalized rat fundus, and normalized rat salivary gland. It was constructed according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the CN1 library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plamid DNA template preparation) comprising: a) a pool of about 34,000 clones from the Rat Unigene Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CA0 and CA1 corresponding to plates R-CA0-AWV through R-CA0-AXS, R-CA0-AZX through R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS, R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA, R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA1-BBA through R-CA1-BDA, R-CA1-BHZ through R-CA1-BJF, R-CA1-BJR, R-CA1-BJT through R-CA1-BKB, R-CA1-BKD, R-CA1-BKF, R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLH through R-CA1-BLN, R-CA1-BLS, R-CA1-BLU-V, R-CA1-BNR, and R-CA1-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CU0s, CW0s, CX0s and normalized libraries CS0, CT0, CU0, CW0, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVV through R-CW0-BWP, R-CW0-BXN through R-CW0-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BV0p (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BFB, R-BV0-ANK through R-BV0-ANR, R-BV0p-AOI through R-BV0p-AOX, and R-BX0-AQY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer

including CY0, CZ0, DA0, DB0, DC0, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYI, R-CZ0-BZB-C, R-DA0-BYJ through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ through R-DB0-BZA, R-DC0-BZI through R-DC0-BZQ, R-DC0-CAY through R-DC0-CBA, R-DD0-BZR through R-DD0-CAA, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CN0, that makes up the tracer. The corresponding plates are R-CN0-BKW through R-CN0-BLD, R-CN0-BLG, R-CN0-BLP through R-CN0-BLR, R-CN0-BLT, R-CN0-BLW-X, R-CN0-BMB, and R-CN0-BMF through R-CN0-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CN0 pool corresponding to the following addresses: bkx-a-09-0-UI, bkx-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI, bkx-d-01-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI, bkx-h-12-0-UI, bky-a-05-0-UI, bkz-a-06-0-UI, bkz-a-11-0-UI, bkz-c-06-0-UI, bkz-c-09-0-UI, bkz-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-e-95-0-UI, bld-1-08-0-UI, bld-f-02-0-UI, blq-h-04-0-UI, blr-a-05-0-UI, blt-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CN0 clone (corresponding to the address bkz-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG_LIB=UI-R-CN1
TAG_TISSUE=cervix
TAG_SEQ=GACCA"

BASE COUNT 177 a 176 c 198 g 211 t
ORIGIN

Query Match 99.1%; Score 173.4; DB 14; Length 762;
Best Local Similarity 99.4%; Pred. No. 2.6e-48;
Matches 174; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTCCAGGAAC TTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 202 AGTCCAGGAAC TTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 143
|||||

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCCCTTCTTTCAGCATGAGAAAGACAAGG 120
|||||
Db 142 AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCCCTTCTTTCAGCATGAGAAAGACAAGG 83
|||||

QY 121 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175
|||||
Db 82 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCCAGCCTAGATAAACACCCAAA 28
|||||

RESULT 2
AI225477
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AI225477
ue88b01.y1 Soares_NMPu Mus musculus cDNA clone linear EST 29-OCT-1998
mRNA sequence.
AI225477
AI225477.1 GI:3808530
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 418)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)

trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES
source
Location/Qualifiers
1..307
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A530061F23"
/clone_lib="RIKEN full-length enriched, adult male aorta and vein"
/sex="male"
/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: Sall; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGAAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 87 a 71 c 70 g 79 t
ORIGIN

Query Match 68.5%; Score 119.8; DB 10; Length 307;
Best Local Similarity 84.0%; Pred. No. 3.7e-30;
Matches 147; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

QY 1 AGTCCAGGAACCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 124 AGTCCAGGAACCTTGACCTTTGTATTGTAAAAATACACATTTCTTAAATGCTCACAAAGC 183
|||||

QY 61 AGGAGGCTCCACACCTCTGGAGGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAGG 120
| |
Db 184 AAG-AGGCTCCACACTCTGCGAGGCCAGGCGCTTCTTTCAGCATGAGAGACAAGG 242
|||||

QY 121 GACAGCAGAGTACTCTCCTCTGGAGGACTAGTCTAGCCCTAGATAAAACACCCAAA 175
|||||

Db 243 AACAGTAGAGTACCTTCCCTCTGGAGGACTGGCCCGTTCTTGAATAAAACACCCAAA 297
|||||

RESULT 6
BB319151
LOCUS BB319151 RIKEN full-length enriched, adult male corpora linear EST 11-JUL-2000
DEFINITION quadrigemina Mus musculus cDNA clone B230378P10 3' similar to AF126063 Mus musculus connective tissue growth factor-like protein precursor (Ctgfl) mRNA, mRNA sequence.
BB319151
ACCESSION BB319151
VERSION BB319151.1 GI:9026186
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 337)
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Watanabe,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watanabe,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:<http://genome.gsc.riken.go.jp/>
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES
source
Location/Qualifiers
1..337
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="B230378P10"
/clone_lib="RIKEN full-length enriched, adult male corpora quadrigemina"
/sex="male"
/tissue_type="corpora quadrigemina"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: Sall; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGAAGGATCCAGAGCTCTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 90 a 88 c 72 g 87 t
ORIGIN

Query Match 68.5%; Score 119.8; DB 10; Length 337;
Best Local Similarity 84.0%; Pred. No. 4e-30;

GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 84 a 82 c 66 g 84 t

ORIGIN

Query Match 64.8%; Score 113.4; DB 10; Length 316;
Best Local Similarity 81.7%; Pred. NO. 6e-28;
Matches 143; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 1 AGTCCAGGAACCTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 133 AGTCCAGGAACCTGACCTTTGTATTTGTATAAATACACTTTTAAATGCTCACAAGC 192
QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCGCTTTCTCTTCAGCATGAGAAAGACAAGG 120
| |||||
Db 193 CAG-AGGCTCCACACTTCTGCCAGGCCAGGCGCTTTTTCAGCATGAGAGACAAGG 251
QY 121 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGAATAAACACCCAAA 175
|||||
Db 252 AACAGTAGAGTACCCTCCTCTGGAGGACTGCGCCGGTCTGGATAAACACCCAAA 306

BB213539 218 bp mRNA linear EST 30-JUN-2000

BB213539 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA clone A530011G24 3' similar to AF126063 Mus musculus connective tissue growth factor-like protein precursor (Ctgfl) mRNA, mRNA sequence.

BB213539 GI:8878492

EST.

house mouse.

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 218)

REFERENCE

AUTHORS

Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Konno,H., et al.)

Unpublished (2000)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/

Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Thermotabilization and thermoactivation of thermostabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1..218

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="A530011G24"

/clone_lib="RIKEN full-length enriched, adult male aorta and vein"

/sex="male"

/tissue_type="aorta and vein"

/dev_stage="adult"

/lab_host="DH10B"

/note="Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGGATCCAGAGCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAATATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 60 a 57 c 45 g 56 t

ORIGIN

Query Match 63.0%; Score 110.2; DB 10; Length 218;
Best Local Similarity 80.6%; Pred. NO. 6e-27;
Matches 141; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 1 AGTCCAGGAACCTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 35 AGTCCAGGAACCTGACCTTTGTATTTGTGCAAAATACACTTTTCTCAACCTCACAACC 94
QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGCGCTTTCTCTTCAGCATGAGAAAGACAAGG 120
| | |||||
Db 95 AAG-AGGCTCCAGCAATTCTGGCAGGCCAGGCGCTTTTCTTCACCATGAGAGACAAGG 153
QY 121 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCCAAA 175
|||||
Db 154 AACAGTAGAGTACCTCTCTGGAGGACTGCCCCGCTTTTGGATAAACACCCAAA 208

RESULT 11

BB186801

LOCUS

DEFINITION

BB186801 RIKEN full-length enriched, adult male spinal cord Mus musculus cDNA clone A330037A13 3' similar to AF126063 Mus musculus connective tissue growth factor-like protein precursor (Ctgfl) mRNA, mRNA sequence.

BB186801

BB186801.1 GI:8847372

EST.

house mouse.

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 289)

REFERENCE

AUTHORS

Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:59:01 ; Search time 1764.82 Seconds
(without alignments)
10471.464 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_635
Perfect score: 635
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	581.4	91.6	1741	10	AF259981	AF259981 Rattus no
2	521.4	82.1	1734	6	AR210324	AR210324 Sequence
3	521.4	82.1	1734	6	AR210325	AR210325 Sequence
4	521.4	82.1	1734	10	AF100778	AF100778 Mus muscu
5	513.6	80.9	1739	10	AF126063	AF126063 Mus muscu
6	273	43.0	1266	6	AX076919	AX076919 Sequence
7	273	43.0	1266	6	AX464186	AX464186 Sequence
8	273	43.0	1283	9	AF083500	AF083500 Homo sapi
9	273	43.0	1293	6	AR210322	AR210322 Sequence
10	273	43.0	1293	6	AR210323	AR210323 Sequence
11	273	43.0	1309	9	AF074604	AF074604 Homo sapi
12	273	43.0	1427	9	AF100780	AF100780 Homo sapi
13	273	43.0	1450	9	BC017782	BC017782 Homo sapi
14	271.4	42.7	841	6	AR210338	AR210338 Sequence
15	261.4	41.2	738	6	AR210337	AR210337 Sequence
16	221.6	34.9	137964	2	AC126895	AC126895 Rattus no
17	221.6	34.9	226303	2	AC095418	AC095418 Rattus no
18	215.4	33.9	137964	2	AC126895	AC126895 Rattus no
19	199.2	31.4	61072	10	AL731698	AL731698 Mouse DNA
20	199.2	31.4	216757	2	AL669906	AL669906 Mus muscu
21	164	25.8	107260	9	AL139352	AL139352 Human DNA
22	94.2	14.8	1987	10	AF218568	AF218568 Rattus no
23	92.8	14.6	1805	5	CHKCEF	J04496 Chicken CEF
24	91	14.3	1871	10	AB015877	AB015877 Rattus no
25	85.6	13.5	1146	6	E13814	E13814 cDNA encodi
26	85.6	13.5	1887	9	AF031385	AF031385 Homo sapi
27	85.6	13.5	1935	9	AF003594	AF003594 Homo sapi
28	85.6	13.5	1985	9	BC001271	BC001271 Homo sapi
29	85.6	13.5	2016	6	AX336386	AX336386 Sequence
30	85.6	13.5	2016	6	AX354335	AX354335 Sequence
31	85.6	13.5	2016	9	HSU62015	U62015 Homo sapien
32	85.6	13.5	2021	9	HSGIG1PRT	Y12084 H.sapiens m
33	85.6	13.5	2025	9	HSCYR61M	Z98053 Homo sapien
34	85.6	13.5	2041	9	BC016952	BC016952 Homo sapi
35	85.6	13.5	2046	9	BC009199	BC009199 Homo sapi
36	85.6	13.5	2052	9	HSCYR61	Y11307 H.sapiens C
37	84.8	13.4	1888	10	AF228049	AF228049 Rattus no
38	84.2	13.3	1743	5	AF320592	AF320592 Xenopus l
39	84	13.2	1418	6	AX206704	AX206704 Sequence
40	83.8	13.2	1480	6	AX206702	AX206702 Sequence
41	83.8	13.2	2018	10	MUSCYR61A	M32490 Mouse Cyr61
42	82.4	13.0	1585	5	NVI271167	AJ271167 Notophtha
43	82.4	13.0	1973	9	HSNOVH	X96584 H.sapiens m
44	82.4	13.0	2270	6	AX035239	AX035239 Sequence
45	82.4	13.0	2270	6	AX342093	AX342093 Sequence

ALIGNMENTS

RESULT 1
AF259981
LOCUS AF259981 1741 bp mRNA linear ROD 09-MAY-2000
DEFINITION Rattus norvegicus CCN family protein COP-1 (Cop-1) mRNA, complete cds.
ACCESSION AF259981
VERSION AF259981.1 GI:7739780
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 1741)
AUTHORS Zhang,R., Averboukh,L., Zhu,W., Zhang,H., Jo,H., Dempsey,P.J.,

Coffey,R.J., Pardee,A.B. and Liang,P.
Identification of rCop-1, a new member of the CCN protein family,
as a negative regulator for cell transformation
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)
98414629
PUBMED
9742130
REFERENCE 2 (bases 1 to 1741)
AUTHORS Liang,P.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer
Center, 649 MRB II, Nashville, TN 37232, USA
FEATURES
-source
1..1741
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1..1741
/gene="Cop-1"
262..1014
/gene="Cop-1"
/note="secreted protein"
/codon_start=1
/product="CCN family protein COP-1"
/protein_id="AAF69011.1"
/db_xref="GI:7739781"
/translation="MRGSLPLRLATSFLLSMVCAQLCRTPCTPWTTPQCPOQVP
LVLDGCGCKVCARRLTESCEHLHVCEPSQGLVCPGAGPGGHGAVCLLDEDDGCEV
NGRRYLDGETFKPNCRLCRDHDGFTCLPLCSEDTVLPWDGPRPKRIQVPGKCCPE
WVCDQGVTPAIQRSAAQHQLSALVTPASADAPWPNWSTAWGPCSTTCGLGIATRVSN
QNRFCLEIQRLRLPRPCLAARSHSSWNSAF"
BASE COUNT 386 a 491 c 480 g 384 t
Query Match 91.6%; Score 581.4; DB 10; Length 1741;
Best Local Similarity 97.8%; Pred. No. 1.4e-129;
Matches 621; Conservative 0; Mismatches 11; Indels 3; Gaps 3;
QY 1 GACGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCC 60
Db 17 GACGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGG-C TTGGCAAGGCTGCAG-C 74
QY 61 GCTGGGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 120
Db 75 GCTGGGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 134
QY 121 CTGTCAGCTTGCTCTAAAGTCTTAGCACTTGTTGGTGGCTTGGGCTTCACACACTGTGAGA 180
Db 135 CTGTCAGCTTGCTCTAAACCCTTAGCACTTGTTGGTGGCTTGGGCTTCACACACTGTGAGA 194
QY 181 CACCTTCGTGGTGGCTCCACGGCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAAGGGAC 240
Db 195 CACCTTCGTGGTGGCTCCACGG-CTCACCTCCAGGTTTGAAGCTGGCTCCACAAAGGGAC 253
QY 241 ACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCTTCCTCTGCCTTC 300
Db 254 ACGGTGACATGAGGGGCAGCCCACTGATCCGCTCTTCTGGCCACTTCCTTCCTCTGCCTTC 313
QY 301 TCTCAATGGTGTGTGCCCCAGCTGTGCCGACACACCCTGTACCTGTCTTGGACACCAACCC 360
Db 314 TCTCAATGGTGTGTGCCCCAGCTGTGCCGACACACCCTGTACCTGTCTTGGACACCAACCC 373
QY 361 AGTGCCACAGGGGGTACCCTGGTGTGGATGGCTGTGGCTGTGCTGCTAAAGTGTGTGCAC 420
Db 374 AGTGCCACAGGGGGTACCCTGGTGTGGATGGCTGTGGCTGTGCTGCTAAAGTGTGTGCAC 433
QY 421 GGAGGCTGGGGAGTCTCTGCGACCACCTGCATGTCTGCGACCCCAAGCCAGGGCCTGGTTT 480
Db 434 GGAGGCTGACGAGTCTCTGCGAACACCTGCATGTCTGCGAACCCAGCCAGGGCCTGGTTT 493
QY 481 GTCAGCCTGGGSCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACG 540
Db 494 GTCAGCCTGGGSCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACG 553
QY 541 GTAGCTGTGAGGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCA 600

Db 554 GTGACTGTGAGGTGAATGGCCGACGGTACCTGGATGGAGAGACCTTTAAACCAATTGCA 613
QY 601 GGGTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635
Db 614 GGGTCCTGTGCGCTGTGATGACGGTGGCTTCACC 648
RESULT 2
AR210324
LOCUS AR210324 1734 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 17 from patent US 6387657.
ACCESSION AR210324
VERSION AR210324.1 GI:21512525
KEYWORDS
SOURCE Unknown..
ORGANISM Unknown..
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A.; Cohen,R.L.; Goddard,A.D.; Gurney,A.L.; Hillan,K.J.;
Lawrence,D.A.; Levine,A.J.; Pennica,D.; Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 17.14-MAY-2002;
FEATURES
source
1..1734
/organism="unknown"
BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN
Query Match 82.1%; Score 521.4; DB 6; Length 1734;
Best Local Similarity 92.0%; Pred. No. 3.8e-115;
Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;
QY 3 CGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGC 62
Db 13 CGCTCCTGATCTCCAGAGGACCCCGGCTGGGACAGGGGCCCTTGGCAGGCTGCAGCTGC 72
QY 63 TG-GGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCC 121
Db 73 TGTGGCAGTAGCTGGGATGGAGGCTTTCTTGTCTGGAACTGAGGAGCTGAGAGGCTCC 132
QY 122 TGTCAG---CTTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCA 178
Db 133 TGTGAGGCTCTTGTCTCTAAACTCTTGGCACTTGGGCTTGGGCTTCACACACTGTCA 192
QY 179 GACACCTTCGTGTGGCCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238
Db 193 GACACCTTCTTGTGGCCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 246
QY 239 ACACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTCTCTCTCTGCCT 298
Db 247 ACACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTCTGCAT 306
QY 299 TCTCTCAATGGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCACC 358
Db 307 TCTCTCAATGGTGTATTCCCACTGTGCCAGCACCCCTGTGCCCTGTCTTGGACACCACC 366
QY 359 CCAGTGCCCAAGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTCTGTAAGTGTGTGC 418
Db 367 CCAGTGCCCAAGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTCTGTCGAGTGTGTGC 426
QY 419 ACGGAGGCTGGGGAGTCTCTGCGACCACCTGCATGTCTGCGACCCCAAGCCAGGCGCTGGT 478
Db 427 ACGGAGGCTGGGGAGTCTCTGCGACCACCTGCATGTCTGCGACCCCAAGCCAGGCGCTGGT 486
QY 479 TTGTCAGCCTGGGGCAGGCCCTGGCGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGA 538
Db 487 TTGTCAGCCTGGGGCAGGCCCTGGCGCCAGTGGCCGTGGTGTGTGTCTCTTGAAGAGGATGA 546
QY 539 CGGTAGCTGTGAGGTGAATGGCCGACGTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598
Db 547 CGGGAGCTGTGAGGTGAATGGCCGACGTACCTGGATGGGAGAGACCTTTAAACCCCAATTG 606

QY 599 CAGGGTCCTGTGCGCGCTGTGATGACGGTGGCTTCACC 635
||||| ||||||| ||||||| ||||||| |||||||
Db 607 CAGGGTTTGTGCGCGCTGTGATGACGGTGGTTTCACC 643
||||| ||||||| ||||||| ||||||| |||||||

RESULT 3
AR210325/c
LOCUS AR210325 1734 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 18 from patent US 6387657.
ACCESSION AR210325
VERSION AR210325.1 GI:215112526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 18 14-MAY-2002;
FEATURES Location/Qualifiers
source
1. .1734
/organism="unknown"

BASE COUNT 393 a 495 c 491 g 355 t
ORIGIN

Query Match 82.1%; Score 521.4; DB 6; Length 1734;
Best Local Similarity 92.0%; Pred. No. 3.8e-115;
Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;

QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGC 62
||||| ||||||| ||||||| ||| ||||||| ||||||| ||||||| ||||||| |||
Db 1722 CGCTCCTGATCTCCAGAGGACCCCGGCTGGGACAGGGGCCCTTGGCGAGGCTGCAGCTGC 1663
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||

QY 63 TG-GGCAGTGGCTTGAATGGAGGTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCC 121
|| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||
Db 1662 TGTGGCAGTAGCTTGGGATGGAGGTCTTTCTTGCTGGGAACCTGAGGAGCTGAGAGGCTCC 1603
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 122 TGTGACG---CTTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCA 178
||||| | |||||| || || |||||| || |||||| || |||||| || |||||| ||
Db 1602 TGTCAGGCTCCTGTCTCTAAACTCTTGGCACTTGGGTGGCTTGGGCTTCACACACTGTCA 1543
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 179 GACACCTTCTGGTGGCCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238
||||||| ||||||||| ||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 1542 GACACCTTCTTGGTGGCCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 1489
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 239 ACACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCTTCTCTGCCT 298
||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 1488 ACACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCATTCTCTTCTCTGCAT 1429
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 299 TCTCTCAATGFTGTGTGCCAGCTGTGCCGGACACCCTGTACCTGTCTCTTGGACACCACC 358
||||||| ||||||| | ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 1428 TCTCTCAATGFTGTATTCCAGCTGTGCCAGACACCCTGTGCCTGTCTCTTGGACACCACC 1369
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 359 CCAGTGCCCCACAGGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGC 418
||||||| ||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 1368 CCAGTGCCCCACGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTGCGAGTGTGTGC 1309
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 419 ACGGAGGCTGGGGAGTCTCTGGACCACTGTCATGTCTGCGACCCAGCCAGGCCCTGGT 478
||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 1308 ACGGAGGCTGGGGAGTCTCTGCGACCACTGTCATGTCTGCGACCCAGCCAGGCCCTGGT 1249
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 479 TTGTACGCTTGGGGCAGGCCCTTGGCGGCATGGGGCTGTGTCTCTTGGATGAGGATGA 538
||||||| ||||||||| ||||| | ||||| || |||||| || ||||| || |||||| ||
Db 1248 TTGTACGCTTGGGGCAGGCCCACTGCGCGTGGTGTGTGTGCCTCTCTGAAGAGGATGA 1189
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 539 CGGTAGCTGTGAGGTGAATGGCCGACGTTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598
||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 1188 CGGGAGCTGTGAGGTGAATGGCCCGCAGGTACCTGGATGGGGAGACCTTTAAACCCCAATTG 1129
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 599 CAGGGTCTCTGCGCGCTGTGATGACGGTGGCTTCACC 635
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||
Db 1128 CAGGGTTTGTGCGCGCTGTGATGACGGTGGTTTCACC 1092
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

RESULT 4
AF100778
LOCUS AF100778 1734 bp mRNA linear ROD 17-DEC-1998
DEFINITION Mus musculus connective tissue growth factor related protein WISP-2 (Wisp2) mRNA, complete cds.
ACCESSION AF100778
VERSION AF100778.1 GI:4028578
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1734)
Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A., Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C., Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D., Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.
WISP genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors
Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)
99061933
9843955

REFERENCE 2 (bases 1 to 1734)
AUTHORS Pennica,D.
Direct Submission
TITLE Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA
JOURNAL Location/Qualifiers
1. .1734
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="C57MG"
/cell_type="epithelial"
/tissue_type="mammary"
/note="transformed by Wnt-1"

1. .1734
/gene="Wisp2"
257. .1012
/gene="Wisp2"
/codon_start=1
/product="connective tissue growth factor related protein WISP-2"
/protein_id="AAC96320.1"
/db_xref="GI:4028579"
/translation="MRGNPLIHLAISFLCILSMVYSQLCPAPCAPCWTPPQPPGVP
LVLDGCGCCRVCAARRLGESCDHLVCDPSSQGLVCQPGAGPSGRGAVCLFEEDDGSCVE
NGRRYLDGETFKPNCRVLCDDGGFTCLPLCSEDSWDGPRPRRIQVPGRCCE
WVCDQAVMQPAIQSSAQGHQLSALVTPASADGCPNWSWTAWPCSTTCGLGIATRV
NQNRFCQLEIQRRLLCLSRPCLASRSHGSWNSAF"

BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN

Query Match 82.1%; Score 521.4; DB 10; Length 1734;
Best Local Similarity 92.0%; Pred. No. 3.8e-115;
Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;

QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGC 62
||||| ||||||| ||||||| ||| ||||||| ||||||| ||||||| ||||||| |||
Db 13 CGCTCCTGATCTCCAGAGGACCCCGGCTGGGACAGGGGCCCTTGGCGAGGCTGCAGCTGC 72
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||

QY 63 TG-GGCAGTGGCTTGAATGGAGGTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCC 121
|| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||
Db 73 TGTGGCAGTAGCTTGGGATGGAGGTCTTTCTTGTGGAACCTGAGGAGCTGAGAGGCTCC 132
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 122 TGTGACG---CTTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCA 178
||||| | |||||| || || |||||| || |||||| || |||||| || |||||| ||
Db 133 TGTGAGGCTCTGTCTCTAAACTCTTGGCACTTGGGTGGCTTGGGCTTCACACACTGTCA 192
||||| |||||| |||||| |||||| |||||| || |||||| || |||||| || |||||| ||

QY 179 GACACCTTCTGTTGGTGGCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238
||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

Db 193 GACACCTTCTTGTCGGCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 246

QY 239 ACACGGTGACATGAGGGGCAGCCCCACTGATCCATCTTCTGGCCACCTTCTCTCTGCCT 298

Db 247 ACACGGTGACATGAGGGGCACCCACTGATCCATCTTCTGGCCATTTCTCTCTGCAT 306

QY 299 TCTCTCAATGTTGTGTCGCCAGCTGTGCCGACACCCCTGTACCTGTCTCTGGACACCACC 358

Db 307 TCTCTCAATGTTGTATTTCCAGCTGTGCCAGCACCCCTGTGCCTGTCTTGGACACCACC 366

QY 359 CCAGTGCCCCACAGGGGGTACCCCTGGTGCTGGATGGCTGTGGCTGCTGTGCTGTGTGC 418

Db 367 CCAGTGCCCCACCGGGGGTACCCCTGGTGCTGGATGGCTGTGGCTGCTGTGAGTGTGTGC 426

QY 419 ACGGAGGCTGGGGAGTCTCTGCGACACCACTGCAATGTCTGCGACCCCGAGGGCCCTGGT 478

Db 427 ACGGAGGCTGGGGAGTCTCTGCGACCACTGCAATGTCTGCGACCCCGAGGGCCCTGGT 486

QY 479 TTGTCAGCCTGGGCGAGGCCCTGGCGGCCCATGGGGCTGTGTGCTCTCTGGATGAGGATGA 538

Db 487 TTGTCAGCCTGGGCGAGGCCCTGGCGGGTGGTGTGTGCTCTCTGGAAGAGGATGA 546

QY 539 CGGTAGCTGTGAGTGAATGGCCGCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598

Db 547 CGGGAGCTGTGAGTGAATGGCCGCGAGGTACCTGGATGGGAGACCTTTAAACCCCAATTG 606

QY 599 CAGGGTCTGTGCGCGCTGTGATGACGGTGGCTTCACC 635

Db 607 CAGGGTTTGTGCGCGCTGTGATGACGGTGGTTTCACC 643

RESULT 5
AF126063

LOCUS
DEFINITION Mus musculus connective tissue growth factor-like protein precursor (Ctgfl) mRNA, complete cds.

ACCESSION AF126063

VERSION AF126063.1 GI:4337059

KEYWORDS

SOURCE Mus musculus.

ORGANISM Mus musculus

REFERENCE
AUTHORS Kumar,S., Hand,A.T., Connor,J.R., Dodds,R.A., Ryan,P.J., Trill,J.J., Fisher,S.M., Nuttall,M.E., Lipshutz,D.B., Zou,C., Hwang,S.M., Votta,B.J., James,I.E., Rieman,D.J., Gowen,M. and Lee,J.C.

TITLE Identification and cloning of a connective tissue growth factor-like cDNA from human osteoblasts encoding a novel regulator of osteoblast functions

JOURNAL J. Biol. Chem. 274 (24), 17123-17131 (1999)

MEDLINE 99287915

PUBMED 10358067

REFERENCE 2 (bases 1 to 1739)

AUTHORS Kumar,S. and Zou,C.

TITLE Direct Submission

JOURNAL Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109, SmithKline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406, USA

FEATURES
source Location/Qualifiers
1..1739
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="lung"
1..1739
/gene="Ctgfl"
242..997
/gene="Ctgfl"
by the sequence presented in GenBank Accession Number AF100778; putative growth factor; CTGF-L; contains IGF binding (IGFBP), Von Willebrand Factor type C (VWC) repeat

and thrombospondin type I (TSPI) domains; member of the CCN (CTGF/Cyr61/Nov) family; lacks the fourth carboxy-terminal (CT) domain present in other members of the CCN family"
/codon_start=1
/product="connective tissue growth factor-like protein precursor"
/protein_id="AAD18058.1"
/db_xref="GI:4337060"
/translation="MRGNPLIHLAIFLCILSMVYSQLCPAPCAPWTPPQCPPGVP LVLDGCGCCRVCAARLGECDHLHVCDP SQGLVCQPGAGPSGRGAVCLFEEDDGSCV NGRYLDGETFKPNCRVLCRDDGFTCLPLCSEDEVRLPSWDGPRRRIQVPGRCPE WVCDOAVMQPAIQPSSAQGHQLSALVTPASADGPCPNWSTAWGPCSTTCGLGIATRV NQNFQLEIQRLCLSRPCLASRSHGSWNSAF"

BASE COUNT 375 a 480 c 489 g 395 t
ORIGIN

Query Match 80.9%; Score 513.6; DB 10; Length 1739;
Best Local Similarity 91.5%; Pred. No. 2.9e-113;
Matches 580; Conservative 0; Mismatches 44; Indels 10; Gaps 3;

QY 6 TTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGCTG- 64

Db 1 TCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCCCTTGGCGAGGCTGCAGCTGCTGT 60

QY 65 GGCAGTGGCTTGAATGGAGGTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCCTGT 124

Db 61 GGCAGTAGCTGGGATGGAGGTCTTTCTTGTGGGAACCTGAGGAGCTGAGAAGCTCCTGT 120

QY 125 CAG---CTTGTCTCTAAAGTCTTAGACACTTGTGGTGGCTTGGGCTTCACACACTGTCAGAC 181

Db 121 CAGGCTCCTGTCTAAACTCTTGGCACCTTGGCGTGGCTTGGGCTTCACACACTGTCAGAC 180

QY 182 ACCTTCGTGTGGCCTCCACGGCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGGACA 241

Db 181 ACCTTCTTGGTGGCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGGACA 234

QY 242 CGGTGACATGAGGGGCAGCCCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCTCTCT 301

Db 235 CGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCATTTCTCTCTCTCTCTCTCT 294

QY 302 CTCAATGGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTTGGACACCCACCCA 361

Db 295 CTCAATGGTGTATTCACAGCTGTGCCCAGCACCCCTGTGCCTGTCTTGGACACCCACCCA 354

QY 362 GTGCCCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAAGTGTGTGCACG 421

Db 355 GTGCCCCACCGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTCGAGTGTGTGCCCG 414

QY 422 GAGGCTGGGGAGTCTCTGCGACCACCTGCATGTCTCGACCCCGAGGCGCTGGTTG 481

Db 415 GAGGCTGGGGAGTCTCTGCGACCACCTGCATGTCTGCGACCCCGAGGCGCTGGTTG 474

QY 482 TCAGCCTGGGCGAGGCCCTTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGG 541

Db 475 TCAGCCTGGGCGAGGCCCTCAGTGGCCGTGTGTGTGTCTCTTCGAAGAGGATGACGG 534

QY 542 TAGCTGTGAGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG 601

Db 535 GAGCTGTGAGTGAACGGCCGCGAGGTACCTGGATGGGAGACCTTTAAACCCCAATTGCAG 594

QY 602 GGTCTGTGCGCGCTGTGATGACGGTGGCTTCACC 635

Db 595 GGTCTGTGCGCGCTGTGATGACGGTGGTTTCACC 628

RESULT 6
AX076919

LOCUS AX076919

DEFINITION Sequence 31 from Patent WO0105836.

ACCESSION AX076919

VERSION AX076919.1 GI:13121575

KEYWORDS

AX076919 1266 bp DNA linear PAT 22-FEB-2001

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1266)
Botstein,D., Goddard,A., Gurney,A.L., Hillan,K.J., Roy,M.A. and Wood,W.I.
TITLE Polypeptidic compositions and methods for the treatment of tumors
JOURNAL Patent: WO 0105836-A 31 25-JAN-2001;
Genentech, Inc. (US)
FEATURES Location/Qualifiers
source 1..1266
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 216 a 418 c 390 g 242 t
ORIGIN
Query Match 43.0%; Score 273; DB 6; Length 1266;
Best Local Similarity 80.9%; Pred. No. 2.5e-55;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCCTTCCCTTCTC 302
DB 4 GGGGACATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTC 63
QY 303 TCAATGGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAG 362
DB 64 TCAAAGGTGCGTACCCAGCTGTGCCGGACACCACTGTACCTGTCCCTGGCCACCTCCCCGA 123
QY 363 TGCCCCACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 422
DB 124 TGCCCCGTGGGAGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 183
QY 423 AGGCTGGGGGAGTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 482
DB 124 TGCCCCGTGGGAGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 183
QY 423 AGGCTGGGGGAGTCCCTGGACCACTGCATGTCTGCGACCCCACTTAAAGTGTGTGACGG 482
DB 184 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCGGCTGGTCTGC 243
QY 483 CAGCCTGGGGCAGGCCCTGGCGGCATGGGCTGTGTCTTGGATGAGGATGACGGT 542
DB 244 CAGCCCGGGCAGGACCGGTGGCGGGCCCTGTATCGGGAAGGGGAGACCTTCCAGCCCACTGCAGC 303
QY 543 AGCTGTGAGGTGAATGCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATTCAGG 602
DB 304 AGCTGTGAGGTGAACGCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCACTGCAGC 363
QY 603 GTCCTGTGCCGCTGTGATGACGGTGGCTTCACC 635
DB 364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACC 396
RESULT 7
AX464186 1266 bp DNA linear PAT 16-JUL-2002
LOCUS Sequence 319 from Patent WO0140466.
DEFINITION AX464186
ACCESSION AX464186
VERSION AX464186.1 GI:21899109
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E., Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L., Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K., Wood,W.L. and Zhang,Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0140466-A 319 07-JUN-2001;
Genentech Inc. (US)
FEATURES Location/Qualifiers
source 1..1266
/organism="Homo sapiens"

/db_xref="taxon:9606"
BASE COUNT 216 a 418 c 390 g 242 t
ORIGIN
Query Match 43.0%; Score 273; DB 6; Length 1266;
Best Local Similarity 80.9%; Pred. No. 2.5e-55;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCCTTCCCTTCTC 302
DB 4 GGGGACATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTC 63
QY 303 TCAATGGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAG 362
DB 64 TCAAAGGTGCGTACCCAGCTGTGCCGGACACCACTGTACCTGTCCCTGGCCACCTCCCCGA 123
QY 363 TGCCCCACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 422
DB 124 TGCCCCGTGGGAGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 183
QY 423 AGGCTGGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCCACTTAAAGTGTGTGTTGT 482
DB 184 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCGGCTGGTCTGC 243
QY 483 CAGCCTGGGGCAGGCCCTGGCGGCATGGGCTGTGTCTTGGATGAGGATGACGGT 542
DB 244 CAGCCCGGGCAGGACCCGCTGGCGGGCCCTGTATCGGGAAGGGGAGACCTTCCAGCCCACTGCAGC 303
QY 543 AGCTGTGAGGTGAATGCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATTCAGG 602
DB 304 AGCTGTGAGGTGAACGCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCACTGCAGC 363
QY 603 GTCCTGTGCCGCTGTGATGACGGTGGCTTCACC 635
DB 364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACC 396
RESULT 8
AF083500 1283 bp mRNA linear PRI 04-NOV-1998
LOCUS Homo sapiens connective tissue growth factor-like protein
DEFINITION precursor, mRNA, complete cds.
ACCESSION AF083500
VERSION AF083500.1 GI:3462835
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1283)
Kumar,S., Hand,A.T., Connor,J.C., Dodds,R.A., Ryan,P.J., Trill,J.G., Fisher,S.M., Slemmon,J.R., Lipshutz,D.B., Bartholomew,V., James,I.E., Rieman,D.J., Gowen,M. and Lee,J.C.
TITLE Identification and cloning of CTGF-L from human osteoblasts, a novel cysteine rich protein containing an IGF binding domain
JOURNAL Bone 23 (5), S240 (1998)
REFERENCE 2 (bases 1 to 1283)
Kumar,S.
AUTHORS Direct Submission
TITLE Submitted (11-AUG-1998) Bone & Cartilage Biology, UW 2109,
JOURNAL SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406,
USA
FEATURES Location/Qualifiers
source 1..1283
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20ql2-ql3"
/cell_type="primary osteoblast"
9..761
/note="CTGF-L; encodes IGF binding (IGFB), Von Willebrand factor type C (VWC) and thrombospondin type I (TSP1) domains; member of the CCN (CTGF/cyr61/nov) family; lacks CDS

the fourth carboxy-terminal domain present in other members of the CCN family"
/codon_start=-1
/product="connective tissue growth factor-like protein precursor"
/protein_id="AAC70350.1"
/db_xref="GI:3462836"
/translation="MRGTPKTHLLAFSLCLLSKVRTQLCPPTCTCPWPPRCPPLGVP
LVLDGGCCRVCAARRLGEPCDQLHVCDAQGLVCPQAGPGGRGALCLLAEDDSSCEV
NGRLYREGETFQPHCSIRCRCEGDFTCVPLCSEDRVLPNSWDCPHPRRVEVLGKCCPE
WVGQGGGLGTQPLPAQGPQFSGLVSSLPGPVPCPEWSTANGSPCTTCGLGMATRVSN
QNRFCRLTQRRLLCLSRPCPPSRGRSPQNSAF"
BASE COUNT 235 a 418 c 389 g 241 t
ORIGIN

Query Match 43.0%; Score 273; DB 9; Length 1283;
Best Local Similarity 80.9%; Pred. No. 2.5e-55;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTCTTGCCCACTTCTCCTCTGCTTCTC 302
Db 3 GGGGACATGAGAGGCACACCGAAGACCCACCTCTGCGCTTCTCCTCTGCTTCTCCTC 62
QY 303 TCAATGGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACCAACCCAG 362
Db 63 TCAAAGGTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGA 122
QY 363 TGCCACAGGGGGTACCCCTGGTGTGGTGTGGTGTGGTGTAAAGTGTGTGCACGG 422
Db 123 TGCCCGTGGGAGTACCCCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTG 182
QY 423 AGGCTGGGGAGTCTCGGACCACTGTCTGCGACCCCACTTCTTGGACACCAACCCAG 482
Db 183 CGGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCGCCAGCGGCTGGTCTGC 242
QY 483 CAGCCTGGGGAGCCCTGGCGGCATGGGGCTGTGTCTTGGATGAGGATGACGGT 542
Db 243 CAGCCCGGGCAGGACCCCGTGGCGGGGCGCTGTGCTTGGCAGAGGACGACAGC 302
QY 543 AGCTGTGAGGTGAATGGCCGAGGTACCTGATGAGGAGACACCTTTAAACCAATGACGG 602
Db 303 AGCTGTGAGGTGAAGCGCGCCTGTATCGGAAGGGGAGACCTTCCAGCCCACTGCAGC 362
QY 603 GTCCTGTGCCGTGTGATGACGGTGGCTTACC 635
Db 363 ATCCGCTGCCGTGCGAGGACGGCGGCTTACC 395

RESULT 9
AR210322
LOCUS AR210322 1293 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 13 from patent US 6387657.
ACCESSION AR210322
VERSION AR210322.1 GI:21512523
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1293)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 13 14-MAY-2002;
FEATURES Location/Qualifiers
source 1. .1293
BASE COUNT 232 a 425 c 393 g 243 t
ORIGIN
Query Match 43.0%; Score 273; DB 6; Length 1293;
Best Local Similarity 80.9%; Pred. No. 2.5e-55;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTCTTGCCCACTTCTCCTCTGCTTCTC 302
Db 16 GGGGACATGAGAGGCACACCGAAGACCCACCTCTGCGCTTCTCCTCTGCTTCTCCTC 75
QY 303 TCAATGGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACCAACCCAG 362
Db 76 TCAAAGGTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGA 135
QY 363 TGCCACAGGGGGTACCCCTGGTGTGGTGTGGTGTGGTGTAAAGTGTGTGCACGG 422
Db 136 TGCCCGCTGGGAGTACCCCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTG 195
QY 423 AGGCTGGGGAGTCTCGGACCACTGTCTGCGACCCCACTTCTTGGACACCAACCCAG 482
Db 196 CGGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCGCCAGCGGCTGGTCTGC 255
QY 483 CAGCCTGGGGAGCCCTGGCGGCATGGGGCTGTGTCTTGGATGAGGATGACGGT 542
Db 256 CAGCCCGGGCAGGACCCCGTGGCGGGGCGCTGTGCTTGGCAGAGGACGACAGC 315
QY 543 AGCTGTGAGGTGAATGGCCGAGGTACCTGATGAGGAGACACCTTTAAACCAATGACGG 602
Db 316 AGCTGTGAGGTGAAGCGCGCCTGTATCGGAAGGGGAGACCTTCCAGCCCACTGCAGC 375
QY 603 GTCCTGTGCCGTGTGATGACGGTGGCTTACC 635
Db 376 ATCCGCTGCCGTGCGAGGACGGCGGCTTACC 408

RESULT 10
AR210323/c
LOCUS AR210323 1293 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 14 from patent US 6387657.
ACCESSION AR210323
VERSION AR210323.1 GI:21512524
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1293)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 14 14-MAY-2002;
FEATURES Location/Qualifiers
source 1. .1293
BASE COUNT 243 a 393 c 425 g 232 t
ORIGIN

Query Match 43.0%; Score 273; DB 6; Length 1293;
Best Local Similarity 80.9%; Pred. No. 2.5e-55;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTCTTGCCCACTTCTCCTCTGCTTCTC 302
Db 1278 GGGGACATGAGAGG@ACACCGAAGACCCACCTCTTGGCTTCTCCTCTGCTTCTCCTC 1219
QY 303 TCAATGGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACCAACCCAG 362
Db 1218 TCAAAGGTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGA 1159
QY 363 TGCCACAGGGGGTACCCCTGGTGTGGTGTGGTGTGGTGTAAAGTGTGTGCACGG 422
Db 1158 TGCCCGCTGGGAGTACCCCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTG 1099
QY 423 AGGCTGGGGAGTCTCGGACCACTGTCTGCGACCCCACTTCTTGGACACCAACCCAG 482
Db 1098 CGGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCGCCAGCGGCTGGTCTGC 1039
QY 483 CAGCCTGGGGAGCCCTGGCGGCATGGGGCTGTGTCTTCTTGGATGAGGATGACGGT 542
Db 1038 CAGCCCGGGCAGGAG@CCCGGTGGCGGGGCGCTTCTTGGCAGAGGACGACAGC 979

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 23, 2003, 09:03:08 ; Search time 168.151 Seconds
(without alignments)
8504.365 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_635
Perfect score: 635
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_101002:*
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	635	100.0	1708	20	AAZ07516
2	521.4	82.1	1734	20	AAZ07517
3	387	60.9	753	20	AAZ07517
4	347	54.6	753	20	AAZ07517
5	318	50.1	681	20	AAZ07521
6	299	47.1	1522	20	AAZ07521
7	278.4	43.8	1337	22	AAH46952
8	278.2	43.8	1352	22	AAH46936
9	273	43.0	1266	21	AAA30048

10	273	43.0	1266	22	AAZ07516	Human CDNA sequenc
11	273	43.0	1266	22	AAZ07516	PRO261 coding sequ
12	273	43.0	1266	22	AAZ07516	Human angiogenesis
13	273	43.0	1267	21	AAZ07516	Human PRO261 cDNA
14	273	43.0	1285	19	AAZ07516	Human connective t
15	273	43.0	1293	20	AAZ07516	Human WISP-2 prote
16	273	43.0	1309	22	AAZ07516	Connective tissue
17	271.4	42.7	841	20	AAZ07516	Human WISP-2 prote
18	268.6	42.3	750	20	AAZ07516	Human WISP-2 prote
19	268.6	42.3	1257	20	AAZ07516	EGF-like homologue
20	261.4	41.2	738	20	AAZ07516	Human WISP-2 prote
21	210	33.1	210	20	AAZ07516	Rat HICP IGFBP dom
22	203.4	32.0	2136	22	AAZ07516	Human full-length
23	164	25.8	13255	22	AAZ07516	Human immune/haema
24	122	19.2	634	22	AAZ07516	Human CDNA 5'-end
25	122	19.2	634	22	AAZ07516	Human CDNA clone r
26	90	14.2	177	20	AAZ07516	Rat HICP VWC domai
27	87.8	13.8	1062	20	AAZ07516	Human connective t
28	87.8	13.8	1062	24	ABK88299	Human connective t
29	85.6	13.5	1146	18	AAZ07516	Human monocyte mat
30	85.6	13.5	1146	24	AAZ07516	Human connective t
31	85.6	13.5	1418	24	ABK48899	DNA encoding human
32	85.6	13.5	1419	24	ABA93130	Human Cyr61 protei
33	85.6	13.5	1887	24	ABQ88131	Human osteoblast d
34	85.6	13.5	2016	22	AAH02896	Human shear stress
35	85.6	13.5	2016	24	ABQ88130	Human osteoblast d
36	85.6	13.5	2016	24	ABK64742	Human benign prost
37	85.6	13.5	2016	24	ABL58558	Kidney cancer rela
38	85.6	13.5	2016	24	ABA93127	Human Cyr61 protei
39	85.6	13.5	2021	24	ABQ88129	Human osteoblast d
40	85.6	13.5	2025	24	ABQ88128	Human osteoblast d
41	85.6	13.5	2052	24	ABQ88127	Human osteoblast d
42	85.6	13.5	2307	21	AAZ07516	Human cancer assoc
43	84	13.2	1418	18	AAZ07516	Human cysteine ric
44	84	13.2	1418	22	AAZ07516	Human cysteine-ric
45	83.8	13.2	1480	22	AAZ07516	Mouse cysteine-ric

ALIGNMENTS

RESULT 1	AAZ07516
ID	AAZ07516 standard; cDNA; 1708 BP.
XX	
AC	AAZ07516;
XX	
DT	26-NOV-1999 (first entry)
XX	
DE	Rat HICP polypeptide encoding cDNA.
XX	
KW	Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW	cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX	
OS	Rattus sp.
XX	
PN	WO9947556-A2.
XX	
PD	23-SEP-1999.
XX	
PF	18-MAR-1999; 99WO-US05999.
XX	
PR	19-MAR-1998; 98US-0044273.
XX	
PA	(TUFT) TUFTS COLLEGE.
XX	
PI	Castellot JJ;
XX	
DR	WPI; 1999-562060/47.
DR	P-PSDB; AAY27434.
XX	
PT	Nucleic acid sequences encoding rat heparin-induced CCN-like protein, -
PT	used in methods to identify modulators or in diagnostic applications -

Claim 2; Fig 1; 108pp; English.

This cdna encodes a rat heparin-induced CCN-like protein (HICP) protein. Agents that stimulate or inhibit HICP protein activity or expression, antisense HICP nucleic acid molecules and HICP antibodies, can be used to modulate cell-associated activity. HICP modulators can be used to treat disorders characterized by aberrant HICP protein activity or expression. probes capable of hybridizing to HICP mRNA or antibodies specific for HICP can be used to detect HICP activity in a biological sample. HICP can be used to treat disorders, such as cardiovascular or fibrotic disorder, characterized by aberrant cell proliferation.

Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
connective tissue growth factor; cancer; melanoma; arteriosclerosis;
leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
connective tissue disorder; catabolic state; inflammation;
testicular-related disorder; angiogenesis; immunological disorder; ss.

Mus sp.

WO9921998-A1.

06-MAY-1999.

29-OCT-1998; 98WO-US22991.

14-APR-1998; 98US-0081695.

29-OCT-1997; 97US-0063704.

03-FEB-1998; 98US-0073612.

(GETH) GENENTECH INC.

Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

WPI; 1999-337420/28.

P-PSDB; AAY17651.

New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

Example 2; Page 178-179, 284pp; English.

The present invention describes Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2 and WISP-3 have homology to connective tissue growth factor (CTGF). Products from the present invention can be used to treat WISP-related disorders such as breast, ovarian, and colon cancer or melanoma. The products can be used to treat arteriosclerosis. The products can also be used to treat other diseases e.g. benign and malignant tumours, leukaemia and lymphoid malignancies, neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal, and blastocoelic disorders, haematopoiesis-related disorders, tissue-growth disorders, skin disorders, desmoplasia, fibrotic lesions, kidney disorders, bone-related disorders such as osteoporosis, trauma such as burns, incisions, and other wounds, connective tissue disorders, catabolic states, testicular-related disorders, and inflammatory, angiogenic and immunologic disorders including arteriosclerosis. The products can also be used for detection and diagnosis especially of individuals with neoplastic cell growth or proliferation. The products can be used in the production of transgenic or knock-out animals. Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing cells.

Sequence 1734 BP; 355 A; 491 C; 495G; 393 T; 0 other;

Db 193 GACACCTTCTTGGTGGCCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAAGG 246

Qy 239 ACACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTCCTTCCTGCGCT 298

Db 247 ACACGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCATTCTCCTCTCGCAT 306

Qy 299 TCCTCAATGGTGTGTGCCCAAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCA 358

Db 307 TCTCTCAATGGTGTATTCCCAAGCTGTGCCCAGCACCCCTGTGCCTGTCTTGGACACCA 366

Qy 359 CCAGTCCCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGCTGTAAAGTGTGTC 418

Db 367 CCAGTCCCCACCGGGGTACCCCTGGTGTGGATGGCTGTGGCTGCTGTGAGTGTGTC 426

Qy 419 ACGGAGCTGGGGAGTCTTGGGACCACTGCATGTCTGGACCCCACTTAAACCCCAATTG 478

Db 427 ACGGAGCTGGGGAGTCTTGGGACCACTGCATGTCTGGACCCCACTTAAACCCCAATTG 486

Qy 479 TTGTACGCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTTGGATGAGGATGA 538

Db 487 TTGTACGCTGGGGCAGGCCCACTGAGTGGCGGTGTGTGTCTTGAAGAGGATGA 546

Qy 539 CGGTAGCTGTGAGGTGAATGCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598

Db 547 CGGAGCTGTGAGGTGAATGCGCGCAGGTACCTGGATGGGAGACCTTTAAACCCCAATTG 606

Qy 599 CAGGCTCTGTGCGCGCTGTGATGACGGTGGCTTCACC 635

Db 607 CAGGCTTTGTGCGCGCTGTGATGACGGTGGTTTCACC 643

RESULT 3

AAZ07517
ID AAZ07517 standard; cDNA; 753 BP.

XX AAZ07517;

DT 26-NOV-1999 (first entry)

XX Rat HICP polypeptide coding sequence.

DE Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;

XX cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.

OS Rattus sp.

PN WO9947556-A2.

XX 23-SEP-1999.

PF 18-MAR-1999; 99WO-US05999.

XX 19-MAR-1998; 98US-0044273.

XX (TUFT) TUFTS COLLEGE.

PA Castellot JJ;

XX WPI; 1999-562060/47.

DR P-PSDB; AAY27434.

XX Nucleic acid sequences encoding rat heparin-induced CCN-like protein, used in methods to identify modulators or in diagnostic applications

PS Claim 5; Fig 1; 108pp; English.

XX The invention provides a rat heparin-induced CCN-like protein (HICP) protein. Agents that stimulate or inhibit HICP protein activity or expression, antisense HICP nucleic acid molecules and HICP antibodies, can be used to modulate cell-associated activity. HICP modulators can be used to treat disorders characterized by aberrant HICP protein activity or expression. Probes capable of hybridizing to HICP mRNA or antibodies

CC specific for HICP can be used to detect HICP activity in a biological sample. HICP can be used to treat disorders, such as a cardiovascular or fibrotic disorder, characterized by aberrant cell proliferation. The present sequence represents the coding sequence of rat HICP.

XX

SQ Sequence 753 BP; 132 A; 235 C; 230 G; 156 T; 0 other;

Query Match 60.9%; Score 387; DB 20; Length 753;
Best Local Similarity 100.0%; Pred. No. 4.6e-100;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 ATGAGGGGAGCCCACTGATCCATCTTGGCCACTTCTGCTCTCTGCTCTCTCAATG 308

Db 1 ATGAGGGGAGCCCACTGATCCATCTTGGCCACTTCTGCTCTCTCTCTCAATG 60

Qy 309 GTGTGTGCCCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCACTGCCCCA 368

Db 61 GTGTGTGCCCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCACTGCCCCA 120

Qy 369 CAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGTAAGTGTGCACGGAGGCTG 428

Db 121 CAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGTAAGTGTGCACGGAGGCTG 180

Qy 429 GGGAGTCTTCCGACCACTGCATGTCTGCCACCCCACTGCTTGGAGGCTGTTGTCAGCCT 488

Db 181 GGGAGTCTTCCGACCACTGCATGTCTGCCACCCCACTGCTTGGAGGCTGTTGTCAGCCT 240

Qy 489 GGGCAGGCCCTGGCGGCCATGGGCTGTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 548

Db 241 GGGCAGGCCCTGGCGGCCATGGGCTGTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 300

Qy 549 GAGTGAATGGCCGAGGTACTGGATGGAGAGACCTTTAAACCCCAATTGACGGGTCTG 608

Db 301 GAGTGAATGGCCGAGGTACTGGATGGAGAGACCTTTAAACCCCAATTGACGGGTCTG 360

Qy 609 TGCGCTGTGATGACGGTGGCTTCACC 635

Db 361 TGCGCTGTGATGACGGTGGCTTCACC 387

RESULT 4

AAZ076489/C
ID AAX76489 standard; DNA; 753 BP.

XX AAX76489;

AC AAX76489;

XX 06-AUG-1999 (first entry)

DT Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18.

DE WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour; connective tissue growth factor; cancer; melanoma; arteriosclerosis; leukaemia; lymphoid malignancy; haematopoiesis-related disorder; tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion; kidney disorder; bone-related disorder; osteoporosis; trauma; burn; connective tissue disorder; catabolic state; inflammation; testicular-related disorder; angiogenesis; immunological disorder; ss.

XX Mus sp.

PN WO9921998-A1.

XX 06-MAY-1999.

PF 29-OCT-1998; 98WO-US22991.

XX 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

XX (GETH) GENENTECH INC.

XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC ocular disorders e.g. corneal infection. The polypeptides can also be
CC used to aid wound healing and epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities. The present
CC sequence represents a human secreted protein encoding cDNA.
XX
SQ Sequence 1337 BP; 257 A; 427 C; 396 G; 252 T; 5 other;

Query Match 43.8%; Score 278.4; DB 22; Length 1337;
Best Local Similarity 78.4%; Pred. No. 4.3e-69;
Matches 345; Conservative 3; Mismatches 84; Indels 8; Gaps 1;

QY 196 CTCACGGCCTCACCTTCAGGTTTGAGCTGGCTCCACAAGGGACACGGTGACATGAGGG 255
Db 2 CTTACAGTTTCACCTTCAGGCTCAAARCTGGSTCTGCA-----GGGGACATGAGAG 53
QY 256 GCAGCCCCACATGATCCATCTTCTGGCCACTTCCTTCCTCTGCCTTCTCTCAATGGTGTG 315
Db 54 GCACACCGAGACCCACCTCCTGGCCCTTCTCCCTCCTCTGCCTCTCTCAAAGGTGCGTA 113
QY 316 CCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTTGGACACACCACCTGCCCCACAGGGGG 375
Db 114 CCCAGCTGTGCCGGACACCATGTACCTGCCCTGGCCACCTCCCCGATGCCCGTGGGAG 173
QY 376 TACCCCTGTGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGGCTGGGGAGT 435
Db 174 TACCCCTGTGCTGGATGGCTGTGGCTGCTGCCGGTATGTGCACGGCGGCTGGGGAGC 233
QY 436 CCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGCCTGTTTGTCTCAGCCTGGGGCAG 495
Db 234 CTTGCGACCACTGCATGTCTGCGACCGCAGCCAGGGCCTGTTTGTCTCAGCCTGGGGCAG 293
QY 496 GCCCTGGCGGCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGA 555
Db 294 GACCCGGTGMCGGGGGGCCCTGTGCCTCTTGGCAGAGGACGACGACGCTGTGAGGTGA 353
QY 556 ATGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTGACGGTCTGTGCCGCT 615
Db 354 ACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCT 413
QY 616 GTGATGACGGTGGCTTCACC 635
Db 414 GCGAGGACGGGGCTTCACC 433

RESULT 8
AAH46936
ID AAH46936 standard; cDNA; 1352 BP.
XX
AC AAH46936;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human secreted protein encoding cDNA (clone Id HBODE48).
XX
KW Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW ophthalmological; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200155430-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01431.
XX
PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
PR 12-SEP-2000; 2000US-0231968.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI Ni J, Ruben SM, Barash SC;
XX
DR WPI; 2001-476220/51.
DR P-PSDB; AAB85526.

XX 17 isolated nucleic acid molecules encoding human secreted proteins,
PT used to preventing, treating or ameliorating a medical condition -
PT
XX
PS Claim 1; Page 415; 482pp; English.

XX The invention provides novel human secreted proteins and polynucleotides
CC encoding them. The secreted proteins can be expressed by standard
CC recombinant methodology. The secreted proteins and polynucleotides are
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can
CC also be used in diagnosing a pathological condition. The antibodies to
CC the proteins can also be used in alleviating symptoms associated with the
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or
CC treated include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC ocular disorders e.g. corneal infection. The polypeptides can also be
CC used to aid wound healing and epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities. The present
CC sequence represents a human secreted protein encoding cDNA.

SQ Sequence 1352 BP; 268 A; 431 C; 399 G; 254 T; 0 other;

Query Match 43.8%; Score 278.2; DB 22; Length 1352;
Best Local Similarity 78.4%; Pred. No. 4.9e-69;
Matches 349; Conservative 0; Mismatches 88; Indels 8; Gaps 1;

QY 191 GTGGCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAGGGACACGGTGACAT 250
Db 3 GTCCGCTTCACAGTTTCACCTTCAGGCTCAAAGCTGGCTCTGCA-----GGGGACAT 54
QY 251 GAGGGGCACGCCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCTCAATGGT 310
Db 55 GAGAGGCACACCGAAGACCCACCTCTGGCCTTCTCCCTCTCTCTCTCTCAAGGT 114
QY 311 GTGTGCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGCCCCACA 370
Db 115 GCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCCGATGCCCGCT 174
QY 371 GGGGTACCCCTGGTGTGGATGGCTGGCTGGCTGTAAAGTGTGTGACGGAGGCTGGG 430
Db 175 GGGAGTACCCCTGGTGTGGATGGCTGGCTGGCTGTCCCGGTATGTGACGGCGGCTGGG 234
QY 431 GGAGTCCCTGGAGACACCTGCATGTCTGCGACCCAGCCAGGGCCTGGTTTGTGAGCCTGG 490
Db 235 GGAGCCCTGGAGCAACTCCACGCTCTGCGACGCCAGCCAGGGCCTGGTCTGCCAGCCCGG 294
QY 491 GGCAGGCCCTGGGGCCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGA 550
Db 295 GGCAGGACCCGGTGGACGGGGGGCCCTGTGCCTCTTGGCAGAGGAGCAGCAGCTGTGA 354
QY 551 GGTGAATGCCCGCAGGTACCTGGATGGAGAGACACCTTTAAACCCAAATTCAGGGTCTCTG 610
Db 355 GGTGAACGGCCGCTGTATCGGGAGGGGAGAGACCTTCCAGCCCCCACTGCAGCATCCGCTG 414

```
QY 611 CCGCTGTGATGACGGTGGCTTCACC 635
Db 415 CCGCTGCGAGGACGGCGGCTTCACC 439

RESULT 9
AAA30048
ID AAA30048 standard; cDNA; 1266 BP.
XX
AC AAA30048;
XX
DT 09-AUG-2000 (first entry)
XX
DE Human PRO261 nucleotide sequence.
XX
KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;
KW PRO317; tumour growth inhibitor; cancer; diagnosis; treatment; human;
KW cell growth; proliferation; growth factor; ADEPT;
KW antibody dependent enzyme mediated prodrug therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200015666-A2.
XX
PD 23-MAR-2000.
XX
PF 08-SEP-1999; 99WO-US20594.
XX
PR 10-SEP-1998; 98US-0099803.
PR 10-SEP-1998; 98WO-US18824.
XX
PA (GETH ) GENENTECH INC.
XX
PI Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI, Botstein D;
XX WPI; 2000-271386/23.
DR P-PSDB; AAY88573.
XX
PT New isolated antibodies which bind to specific polypeptides used for
PT diagnosis and treatment of neoplastic cell growth and proliferation -
XX
PS Example 7; Fig 13; 200pp; English.
XX
CC This sequence represents a human PRO261 nucleotide sequence. PRO261 is a
CC growth factor. The invention relates to isolated antibodies which bind to
CC a polypeptide. The "PRO" polypeptides are encoded by genes which are over
CC expressed in the genome of tumour cells. Vectors and host cells
CC comprising the nucleic acid encoding the antibodies are used in the
CC production of the antibodies. The antibodies and nucleic acids encoding
CC them are used for diagnosing a tumour in a mammal. The antibodies are
CC used for inhibiting the growth of tumour cells and identifying compounds
CC that inhibit a biological or immunological activity of and/or expression
CC of a PRO187, PRO533, PRO214, PRO240, PRO211, PRO230, PRO261, PRO246 or
CC PRO317 polypeptide. The antibody can be used in antibody dependent enzyme
CC mediated prodrug therapy (ADEPT) by conjugating the antibody to a
CC prodrug-activating enzyme which converts a prodrug to an anti-cancer
CC drug. The antibodies can be fluorescently labelled and monitored by light
CC microscopy, flow cytometry or fluorimetry for diagnosis and prognosis of
CC tumours.
XX
SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 43.0%; Score 273; DB 21; Length 1266;
Best Local Similarity 80.9%; Pred. No. 1.4e-67;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCTTCCTGCTTC 302
Db 4 GGGACATGAGAGGCACACCGAAGACCCCACTCCTGGCCTTCTCCCTCCTCTGCTTC 63

QY 303 TCAATGGTGTGCCCCAGCTGTGCCGGACACCCTGTACCTGTCTTGGACACACCCAC 362
Db 64 TCAAAGGTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCCCTGGCCACCTCCCCGA 123
```

```
QY 363 TGCCACACAGGGGTACCCCTGGTGGCTGGATGGCTGTGGCTGCTGTAAGTGTGTGCACGG 422
Db 124 TGCCCGCTGGGAGTACCCCTGGTGGCTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG 183

QY 423 AGGCTGGGGAGTCTCTGCGACCACCTGCATGCTCTCGACCCCAAGGGCCCTGGTTGT 482
Db 184 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGC 243

QY 483 CAGCCTGGGGCAGGCCCTGGGGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT 542
Db 244 CAGCCCGGGCAGGACCCGGTGGCCGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGC 303

QY 543 AGCTGTAGGTGAATGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAATTGCAGG 602
Db 304 AGCTGTAGGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCCACTGCAGC 363

QY 603 GTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635
Db 364 ATCCGCTGCGCTGCGAGGACGGCGGCTTCACC 396

RESULT 10
AAS21403
ID AAS21403 standard; cDNA; 1266 BP.
XX
AC AAS21403;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human cDNA sequence encoding for PRO261 polypeptide.
XX
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200140466-A2.
XX
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000WO-US32678.
XX
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.
XX
```


(GETH) GENENTECH INC.

Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W; Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S; Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

WPI; 2001-408281/43.
P-PSDB; AAU12331.

Isolated, secretory and transmembrane PRO polypeptide used to detect other PRO polypeptides, link bioactive molecules to cells expressing PRO polypeptides, and detect the presence of mammalian tumours e.g. lung, breast, prostate, cervical -

Claim 3; Fig 319; 813pp; English.

AAS21244-AAS21518 encode for novel human secretory and transmembrane PRO polypeptides. The PRO polypeptides are useful to detect other PRO polypeptides, to link bioactive molecules to cells expressing PRO polypeptides, to modulate biological activities of cells expressing PRO polypeptides, and to detect the presence of mammalian lung, colon, breast, prostate, rectal, cervical or liver tumours by comparing PRO polypeptide expression in a cell sample to that in a control sample. Some of the 275 sequences are also useful to stimulate the release of tumour necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or differentiation of chondrocytes, the proliferation or gene expression in pericyte cells, the release of proteoglycans from cartilage, the proliferation of inner ear utricular supporting cells or of T-lymphocytes, the release of a cytokine from peripheral blood monocytes (PBMCs), or the proliferation of endothelial cells. Some of the PRO polypeptides may modulate glucose or free fatty acid uptake by skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide to factor VIIA. The PRO polypeptides can be used in assays to identify molecules involved in binding interactions. The polynucleotides encoding PRO polypeptides can be used to generate probes, antisense RNA/DNA, transgenic or knock out animals and can be used in gene therapy.

Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 43.0%; Score 273; DB 22; Length 1266;
Best Local Similarity 80.9%; Pred. No. 1.4e-67;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY	243	GGTGACATGAGGGGACGCCCACTGATCCATCTTCGGCCACTTCCTTCCCTCTGCCTTC	302
Db	4	GGGACATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTC	63
QY	303	TCAATGGTGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCCACCCAG	362
Db	64	TCAAAGGTGCTATCCACAGCTGTGCCCGACACCATGTACCTGCCCTTGGCCACCTCCCCGA	123
QY	363	TGCCCCACAGGGGTACCCCTGGTGGTGGATGGCTGGCTGGCTGTAAAGTGTGTGCACGG	422
Db	124	TGCCCCGTGGGAGTACCCCTGGTGGTGGATGGCTGTGGCTGTGCGGGTGTATGTGCACGG	183
QY	423	AGGCTGGGGAGTCTCGACCACTGCATGTCTGGACCCCGAGCCAGGCGCTGGTTGT	482
Db	184	CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCCAGGCGCTGGTCTGC	243
QY	483	CAGCCTGGGGCAGGCCCTTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT	542
Db	244	CAGCCCGGGCAGGACCCGGTGGCCGGGGGGCCCTGTGCTCTTGGCCTCTTGGCAGAGGACGACG	303
QY	543	AGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG	602
Db	304	AGCTGTGAGGTGAACGGCCGCCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGC	363
QY	603	GTCCTGTGCCGTGTGATGACGGTGGCTTACC	635
Db	364	ATCCGCTGCCGTGCGAGGACGGCGGCTTACC	396

```

AAF60368
ID   AAF60368 standard; cDNA; 1266 BP.
XX
AC   AAF60368;
XX
DT   27-APR-2001 (first entry)
XX
DE   PRO261 coding sequence.
XX
KW   Cytostatic; PRO protein; tumour; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200105836-A1.
XX
PD   25-JAN-2001.
XX
PF   20-DEC-1999; 99WO-US30999.
XX
PR   20-JUL-1999; 99US-0144758.
PR   26-JUL-1999; 99US-0145698.
PR   08-SEP-1999; 99WO-US20594.
PR   13-SEP-1999; 99WO-US20944.
PR   15-SEP-1999; 99WO-US21090.
PR   05-OCT-1999; 99WO-US23089.
PR   29-NOV-1999; 99WO-US28214.
PR   30-NOV-1999; 99WO-US28313.
PR   02-DEC-1999; 99WO-US28564.
XX
( GETH ) GENENTECH INC.
PA
XX
PI   Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;
XX
DR   WPI; 2001-091968/10.
DR   P-PSDB; AAB68598.
XX
PT   New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,
PT   useful for diagnosing and treating cancers -
XX
PS   Claim 50; Fig 13; 196pp; English.
XX
CC   The present invention relates to PRO proteins and coding sequences. The
CC   present sequence is the coding sequence for one such PRO protein.
CC   It was found that the PRO genes are amplified in the genome of tumour
CC   cells. The gene amplification is expected to be associated with the
CC   overexpression of the gene product and contributes to tumourigenesis.
CC   Therefore, antagonists of PRO proteins are useful for the treatment of
CC   benign or malignant tumours, leukaemias, lymphoid malignancies and other
CC   disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
CC   epithelial, inflammatory and immunologic disorders.
XX
SQ   Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

    Query Match      43.0%; Score 273; DB 22; Length 1266;
    Best Local Similarity 80.9%; Pred. No. 1.4e-67;
    Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY   243 GGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCCTTCCCTCTGCCCTTC 302
Db   4 GGGGACATGAGAGGACACACCGAAGACCCACCTCCTGGCCCTTCTCCCTCCTCTGCCCTC 63

QY   303 TCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCCTTGGACACCCACCCAG 362
Db   64 TCAAAGTTCGTAGCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCCGA 123

QY   363 TGCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGG 422
Db   124 TGCCCGCTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG 183

QY   423 AGGCTGGGGAGTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGCGCTGGTTGT 482
Db   184 CGGCTGGGGAGCCCTGCGACCAACTCCACGCTCTGCGACGCCAGCCAGGCGCTGGTCTGC 243

```


Search completed: July 23, 2003, 11:15:17
Job time : 169.151 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:55:06 ; Search time 34.5474 Seconds
(without alignments)
5636.882 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_635
Perfect score: 635
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GTGATGACGGTGGGCTTCACC 635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	521.4	82.1	1734	4	US-09-182-145-17
c 2	521.4	82.1	1734	4	US-09-182-145-18
3	273	43.0	1293	4	US-09-182-145-13
c 4	273	43.0	1293	4	US-09-182-145-14
5	271.4	42.7	841	4	US-09-182-145-39
6	261.4	41.2	738	4	US-09-182-145-38
7	87.8	13.8	1062	4	US-09-253-316-3
8	84	13.2	1418	4	US-09-142-569-3
9	83.8	13.2	1480	4	US-09-142-569-1
10	79.2	12.5	1128	2	US-08-459-101A-1
11	78	12.3	1766	4	US-09-182-145-9
c 12	78	12.3	1766	4	US-09-182-145-10
13	75.2	11.8	2830	4	US-09-182-145-1
c 14	75.2	11.8	2830	4	US-09-182-145-2
15	69.8	11.0	1142	4	US-09-253-316-1
16	69.8	11.0	1212	4	US-09-182-145-34
c 17	69.8	11.0	1212	4	US-09-182-145-35
18	69.8	11.0	1335	4	US-09-182-145-30
c 19	69.8	11.0	1335	4	US-09-182-145-31
20	66	10.4	2075	1	US-08-167-628-1
21	66	10.4	2075	1	US-08-386-680-1
22	66	10.4	2075	1	US-08-459-717-1
23	66	10.4	2075	1	US-08-712-302-1
24	66	10.4	2075	2	US-08-880-031-1
25	66	10.4	2075	3	US-09-097-179-1
26	66	10.4	2075	4	US-09-080-715-1
27	66	10.4	2075	4	US-09-142-569-7

28	66	10.4	2075	5	PCT-US96-08140-1	Sequence 1, Appli
29	66	10.4	2998	3	US-09-054-368-1	Sequence 1, Appli
30	66	10.4	2998	3	US-09-054-274-1	Sequence 1, Appli
31	66	10.4	2998	4	US-09-056-704-1	Sequence 1, Appli
32	63.4	10.0	693	4	US-09-182-145-24	Sequence 24, Appli
33	63.4	10.0	1101	4	US-09-182-145-29	Sequence 29, Appli
34	63.4	10.0	1202	4	US-09-182-145-26	Sequence 26, Appli
35	63.4	10.0	1403	4	US-09-182-145-23	Sequence 23, Appli
36	62.6	9.9	2350	4	US-09-187-478-1	Sequence 1, Appli
37	62.6	9.9	2350	4	US-09-292-036-1	Sequence 1, Appli
38	60.8	9.6	2267	4	US-09-142-569-5	Sequence 5, Appli
39	55.8	8.8	4214	4	US-09-122-135-1	Sequence 1, Appli
40	48.6	7.7	2541	2	US-08-656-393-1	Sequence 1, Appli
41	41.4	6.5	51	4	US-09-182-145-117	Sequence 117, App
c 42	36.6	5.8	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c 43	36.6	5.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
44	35.2	5.5	2300	3	US-08-486-343A-2	Sequence 2, Appli
45	34.4	5.4	87350	3	US-08-781-891-79	Sequence 79, Appli

ALIGNMENTS

RESULT 1
US-09-182-145-17
; Sequence 17, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP. POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-182-145-17

Query Match	82.1%;	Score	521.4;	DB	4;	Length	1734;
Best Local Similarity	92.0%;	Pred. No.	1.1e-135;				
Matches	586;	Conservative	0;	Mismatches	41;	Indels	10;
Gaps	3;						
QY	3	CGCTTCTGATCTCCAGAGGACCCTGGGGTGGACAGGGCCCTTGGCAAGGCTGCAGCCGC	62				
Db	13	CGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGCCCTTGGCGAGGCTGCAGCTGC	72				
QY	63	TG-GGCAGTGGCTTGGAAATGGAGGCTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCC	121				
Db	73	TGTGGCAGTAGCTTGGGATGGAGGCTCTTTCTGTGGAACTGAGGAGCTGAGAGGCTCC	132				
QY	122	TGTCAG---CTTGTCCTAAAGTCTTAGCAGTGTGGTGGCTTGGGCTTCACACACTGTCA	178				
Db	133	TGTCAGGCTCCTCTCCTAAACTCTTGGCAGCTTGGGTGGCTTGGGCTTCACACACTGTCA	192				
QY	179	GACACCTTCGTGTGGTGGCCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG	238				

Db 193 GACACCTTCTTGGTGGCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 246

QY 239 ACACGGTGACATGAGGGGACAGCCCACTGATCCATCTTCTGGCCACTTCCCTCCTCGCCT 298

Db 247 ACACGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCATTTCCCTCCTCGCAT 306

QY 299 TCTCTCAATGGTGTGCTGCCCCAGCTGTGCGGACACCCCTGTACCTGTCTTGAAGTGTGTGC 358

Db 307 TCTCTCAATGGTGTATTTCCAGCTGTGCCCCAGACCCCTGTGCCTGTCTTGGACACCACC 366

QY 359 CCAGTGCCCCACAGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGTCTTGAAGTGTGTGC 418

Db 367 CCAGTGCCCCACCGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGTCTCGAGTGTGTGC 426

QY 419 ACGGAGGCTGGGGGAGTCTCTGCGACCACTGTCATGTCTGCGACCCCAAGCCAGGGCCTGGT 478

Db 427 ACGGAGGCTGGGGGAGTCTCTGCGACCACTGTCATGTCTGCGACCCCAAGCCAGGGCCTGGT 486

QY 479 TTGTCAGCCTGGGGGAGTCTCTGCGACCACTGTCATGTCTGCGACCCCAAGCCAGGGCCTGGT 538

Db 487 TTGTCAGCCTGGGGGAGTCTCTGCGACCACTGTCATGTCTGCGACCCCAAGCCAGGGCCTGGT 546

QY 539 CGGTAGCTGTGAGTGAATGGCCGAGTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598

Db 547 CGGGAGCTGTGAGTGAATGGCCGAGTACCTGGATGGGAGACCTTTAAACCCCAATTG 606

QY 599 CAGGGTCTCTGTCGCTGTGATGACGGTGGCTTCACC 635

Db 607 CAGGGTCTCTGTCGCTGTGATGACGGTGGCTTCACC 643

RESULT 2

US-09-182-145-18/c

; Sequence 18, Application. US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P1176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 18

; LENGTH: 1734

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-182-145-18

Query Match 82.1%; Score 521.4; DB 4; Length 1734;

Best Local Similarity 92.0%; Pred. No. 1.le-135;

Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;

QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGTGGGACAGGGCCCTTGGCAGGCTGCAGCCGC 62

Db 1722 CGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGCCCTTGGCGAGGCTGCAGCTGC 1663

QY 63 TG-GGCAGTGGCTGGAATGGAGGCTCTTATTACTGGGAACTGAGGAGCTAAGAGGCTCC 121

Db 1662 TGTGGCAGTAGGTTGGGATGGAGGTCTTTCTTCTGGGAACTGAGGAGCTGAGAGGCTCC 1603

QY 122 TGTGAG---CTTGTCTTAAAGTCTTAGCAGCTTGTGGTGGCTTGGGCTTACACACACTGTCA 178

Db 1602 TGTGAGGCTCCTGTCTTAAACTCTTGGCAGCTTGGCGTGGCTTGGGCTTACACACACTGTCA 1543

QY 179 GACACCTTCTGTTGGCCTTCCACGGCCTCACCTTACAGTTTGAAGCTGGCTCCACAAAGGG 238

Db 1542 GACACCTTCTTGGTGGCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAAGGG 1489

QY 239 ACACGGTGACATGAGGGGACAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTCTGCCT 298

Db 1488 ACACGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCATTTCTTCTCTCTGCAT 1429

QY 299 TCTCTCAATGGTGTGCTGCCAGCTGTGCCGAGACACCCCTGTACCTGTCTCTTGGACACCACC 358

Db 1428 TCTCTCAATGGTGTATTTCCAGCTGTGCCAGCACCCTGTGCTTGTCCCTTGGACACCACC 1369

QY 359 CCAGTGCCCCACAGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGTCTTAAAGTGTGTGC 418

Db 1368 CCAGTGCCCCACAGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGTCTGAGTGTGTGC 1309

QY 419 ACGGAGGCTGGGGAGTCTCTGCGACCACTGCGACCCCAAGCTGTCTGCGACCCCAAGGCTGGT 478

Db 1308 ACGGAGGCTGGGGAGTCTCTGCGACCACTGCGACCCCAAGCTGTCTGCGACCCCAAGGCTGGT 1249

QY 479 TTGTCAGCCTGGGGCAGGCCCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGA 538

Db 1248 TTGTCAGCCTGGGGCAGGCCCCCAGTGGCCGCTGTGTGCTGTGTGCTCTCTCGAAGAGGATGA 1189

QY 539 CGGTAGCTGTGAGTGAATGGCCGAGTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598

Db 1188 CGGGAGCTGTGAGTGAATGGCCGAGTACCTGGATGGGAGACCTTTAAACCCCAATTG 1129

QY 599 CAGGGTCTCTGTCGCTGTGATGACGGTGGCTTCACC 635

Db 1128 CAGGGTCTCTGTCGCTGTGATGACGGTGGCTTCACC 1092

RESULT 3

US-09-182-145-13

; Sequence 13, Application. US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P1176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 13

; LENGTH: 1293

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-182-145-13

Query Match 49.0%; Score 273; DB 4; Length 1293;

Best Local Similarity 80.9%; Pred. No. 1e-66;

Db 126 TGCCCGCTGGAGTACCCCTGGTGGTGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG 185
QY 423 AGGCTGGGGAGTCTCGACACCACTGCATGTCTCGACCCCGACGAGGGCCCTGGTTTGT 482
Db 186 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTCGACGCCAGCGCCCTGGTCTGC 245
QY 483 CAGCCTGGGACGGCCCTGGCGGCCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGT 542
Db 246 CAGCCCGGGGAGGACCCGGTGGCCGGGGGCCCTGTGCTCTTGGCAGAGGACGACAGC 305
QY 543 AGCTGTAGGTGAATGGCCGACGCTACCTGGATGGAGAGACCTTTAAACCCAATTGCAGG 602
Db 306 AGCTGTAGGTGAACGGCCGCTGTATCGGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 365
QY 603 GTCCTGTGCCGCTGTGATGACGGTGGCTTCAACC 635
Db 366 ATCCGCTGCCGCTGCGAGGACGGCGGCTTCAACC 398
RESULT 6
US-09-182-145-38
; Sequence 38, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 38
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-38

Query Match 41.2%; Score 261.4; DB 4; Length 738;
Best Local Similarity 81.1%; Pred. No. 1.4e-63;
Matches 304; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 261 CCACTGATCCATCTTCTGGCCACTTCTCCTCCTCTGCTTCTCTCAATGGTGTGCCAG 320
Db 1 CCGAAGACCCACCTCTCTGGCCTTCTCCTCCTCTGCTTCTCTCAAGGTGCGTACCCAG 60
QY 321 CTGTGCCGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCACAGGGGTACCC 380
Db 61 CTGTGCCGACACCACTGTACCTGCCCTGGCCACCTCCCGATGCCCGTGGAGTACCC 120
QY 381 CTGGTGTGGATGGCTGTGGCTGTCTGTAAAGTGTGTGCACGGAGCTGGGGAGTCTCTGC 440
Db 121 CTGGTGTGGATGGCTGTGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 441 GACCACCTGCGATGTCTGCGACCCAGCAGGGCCCTGGTTTGTACGCTGGGGCAGGCCCT 500
Db 181 GACCAACTCCACGTCTGCGACGCCAGCAGGGCCCTGGTCTGCCAGCCCCGGGACGACCC 240
QY 501 GCGGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAATGCC 560

Db 241 GGTGGCGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGAGGTGAGGTGAACGGC 300
QY 561 CGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTGTGCCGCTGTGAT 620
Db 301 CGCCTGTATCGGGAAGGGAGACCTTCCAGCCCCCACTGCAGCATCCGCTGCCGCTGCCGAG 360
QY 621 GACGGTGGCTTCAACC 635
Db 361 GACGGCGGCTTCAACC 375
RESULT 7
US-09-253-316-3
; Sequence 3, Application US/09253316
; Patent No. 6395890
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Jaspers, Stephen R.
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR HOMOLOGS
; FILE REFERENCE: 97-75
; CURRENT APPLICATION NUMBER: US/09/253,316
; CURRENT FILING DATE: 1999-02-19
; EARLIER APPLICATION NUMBER: US 60/075,300
; EARLIER FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate sequence
; NAME/KEY: misc_feature
; LOCATION: (1)...(1062)
; OTHER INFORMATION: n = A,T,C or G
US-09-253-316-3

Query Match 13.8%; Score 87.8; DB 4; Length 1062;
Best Local Similarity 36.8%; Pred. No. 2.7e-15;
Matches 114; Conservative 53; Mismatches 143; Indels 0; Gaps 0;
QY 318 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCACAGGGGTA 377
Db 136 CARTTYGYCAYTGGCCNTGYAARTGYCCNCARCARAARCCNMGNMTGYCCNCCNGNGTN 195
QY 378 CCCCTGGTGGATGGCTGTGGCTGTGTAAAGTGTGTGCACGGAGGCTGGGGAGTCC 437
Db 196 WSNYTNMGNGAYGGNTGYGGNTGYTGYAARATHHTGYGCNAAARCCNCGNGARATH 255
QY 438 TGCACACCTGCATGTCTGCGACCCAGCCAGGGCCCTGGTTTGTACGCTGGGGCAGGC 497
Db 256 TGYAAYGARGCNGAYTNTGYGAYCCNCAAYAARGGNYTNTAYTGYGAYTAYWSNGTNAY 315
QY 498 CCTGGCGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGTPAGCTGTGAGGTGAAT 557
Db 316 MGNCCNMGNATAYGARACNGGNGTNTGYCNCNTAYTNGTNGCNGNGGNTGYGARTTYAAY 375
QY 558 GGCCGCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTGTGCCGCTGT 617
Db 376 CARGTNCAYTAYCAYAAAYGGNCARGTNTTYCARCCNAAAYCCNYTNTYWSNTGYTNTGY 435
QY 618 GATGACGGTG 627
Db 436 GTNWSNGNG 445
RESULT 8
US-09-142-569-3
; Sequence 3, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:

APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,569
; FILING DATE: 02-Apr-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 124..1266
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "Human cyr61 cdna coding sequence"
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-142-569-3

Query Match 13.2%; Score 84; DB 4; Length 1418;
Best Local Similarity 54.5%; Pred. No. 3.4e-14;
Matches 168; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 324 TGCCGGACACCCTGTACCTGTCTTGGACACCAACCCAGTGCCTCCACAGGGGGTACCCCTG 383
Db 199 TGCCCCGGCTGCCTGCCACTGCCCCCTGGAGCGCCCAAGTGCCTCCGCGGGAGTCGGGCTG 258
QY 384 GTGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGGCTGGGGGAGTCTCTGCGAC 443
Db 259 GTCCGGGACGGCTGCGGCTGCTGTAAAGTGTGTGCACGGAGGCTCAACGAGGACTGCAGC 318
QY 444 CACCTGCATGCTGTGCGACCCAGCCAGGGCTGGTTTGTACGCTGGGGCAGGCCCTGGC 503
Db 319 AAAACGCAGCCCTGCGACCAACCAAGGGCTGGAATGCAACTTCGCGCGCCAGCTCCACC 378
QY 504 GGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGC 563
Db 379 GCTCTGAAGGGGATCTGCAGAGCTCAGTCAGAGGGCAGACCCCTGTGATATATACTCCAGA 438
QY 564 AGGTACCTGGATGGAGAGACCTTTAAACCAATTCAGGGTCTCTGCGCGCTGTGATGAC 623
Db 439 ATCTACCAACACGGGGAAAGTTTCCAGCCCAACTGTCAACATCAGTGCACATGTATTGAT 498
QY 624 GGTGGCTT 631
Db 499 GCGCCGT 506

RESULT 9

US-09-142-569-1
; Sequence 1, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:
; APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,569
; FILING DATE: 02-Apr-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 180..1316
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "Mouse cyr61 cdna coding sequence"
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-142-569-1

Query Match 13.2%; Score 83.8; DB 4; Length 1480;
Best Local Similarity 51.2%; Pred. No. 3.9e-14;
Matches 196; Conservative 0; Mismatches 187; Indels 0; Gaps 0;
QY 249 ATGAGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCCTTCCCTCTGCTTCTCAATG 308
Db 180 ATGAGCTCCAGCACCTTCAGGACGCTCGCTGTCGCGCTCACCCCTTCTCCACTTGACCAGA 239
QY 309 GTGTGTCCCACTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCCCA 368
Db 240 CTGGCGCTCTCCACCTGCCCGCGCGCTGCCACTGCCCTCTGGAGGCACCAAGTGGCC 299
QY 369 CAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGCACGGAGGCTG 428
Db 300 CCGGGAGTCGGGTGGTCCGGGACGGCTGCGGCTGCTGTAAAGTCTGCGCTAAACAACCTC 359
QY 429 GGGGAGTCTCGGACCACTGCATGTCTGCGACCCCAACCCAGCCCTGGTGTGTGTCAGCCT 488
Db 360 AACGAGGACTGCAGCAAAACTCAGCCCTGCGACCAACCAAGGGTGTGAATGCAATTTC 419
QY 489 GGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 548
Db 420 GCGCCAGCTCCACCGCTCTGAAAGGATCTGCAGAGCTCAGTCAGAGGAGGACACCCCTGT 479
QY 549 GAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTCAGGGTCTCTG 608

Db 480 GAATATAACTCAGAAATCTACCAAAACGGGAAAGCTTCCAGCCCCCAACTGTAAACACCAG 539

QY 609 TGCCGCTGTGATGACGGTGGCTT 631

Db 540 TGCACATGTATTGATGGCCCGT 562

RESULT 10

US-08-459-101A-1

; Sequence 1, Application US/08459101A

; Patent No. 5945300

; GENERAL INFORMATION:

; APPLICANT: LI, ET AL.

; TITLE OF INVENTION: Connective Tissue Growth Factor-2

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

; ADDRESSEE: CECCHI, STEWART & OLSTEIN

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,101A

; FILING DATE: June 2, 1995

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/07736

; FILING DATE: 12 JUL 94

; ATTORNEY/AGENT INFORMATION:

; NAME: FERRARO, GREGORY D.

; REGISTRATION NUMBER: 36,134

; REFERENCE/DOCKET NUMBER: 325800-317

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1128 BASE PAIRS

; TYPE: NUCLEIC ACID

; STRANDEDNESS: SINGLE

; TOPOLOGY: LINEAR

; MOLECULE TYPE: cDNA

US-08-459-101A-1

Query Match 12.5%; Score 79.2; DB 2; Length 1128;

Best Local Similarity 52.1%; Pred. No. 6.7e-13;

Matches 177; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 290 CCTCTGCCTTCTCTCAATGGTGTGTCGCCAGCTGTGCGGACACCCCTGTACCTGTCTTG 349

Db 42 CCTTCTCACTTGACCAGGTGGGGCTCTCCACCTGCCCGCTGACTGCCACTGCCCCCT 101

QY 350 GACACCACCCAGTGCCACAGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTGTA 409

Db 102 GGAGGCGCCCAAGTGCGCCGGGAGTGGGCTGGTCCGGACGGCTGCGGCTGTTGTA 161

QY 410 AGTGTGTGCACGGAGGCTGGGGGAGTCCCTGCGACCACTGCATGTCTGCGACCCCA 469

Db 162 GGTCTGCGCCCAAGCAGCTCAACGAGGACTGCAGAAAACGCAGCCCTGCGACCA 221

QY 470 GGGCCTGGTTGTACGCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGA 529

Db 222 GGGGCTGGAATGCAACTTCGGCGCCAGCTCCACCGCTCTGAAGGGGATCTGCAGAGCTCA 281

QY 530 TGAGGATGACGGTAGCTGTGAGGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAA 589

Db 282 GTCAGAGGGCAGACCCTGTGAATATAACTCCAGAATCTACCAAAACGGGAAAGTTTCCA 341

QY 590 ACCCAATTGCAGGGTCTCTGTGCCGCTGTGATGACGGTGGC 629

Db 342 GCCCAACTGTAAACATCAGTGCACATGTATTGGATGGCGC 381

RESULT 11

US-09-182-145-9

; Sequence 9, Application US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P1176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 9

; LENGTH: 1766

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: Unsure

; LOCATION: 10

; OTHER INFORMATION: Unknown base

US-09-182-145-9

Query Match 12.3%; Score 78; DB 4; Length 1766;

Best Local Similarity 58.2%; Pred. No. 1.7e-12;

Matches 165; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 316 CCCAGCTGTGCCGACACCCCTGTACCTGTCTCTTGGACACCCACCCAGTGCACAGGGG 375

Db 215 CCGAATTCGCAAGTGGCCATGTGAGTGGCCACAATCCCCACCTCGCTGCCACTGGGCG 274

QY 376 TACCCCTGGTCTGGATGGCTGTGGCTGTAAAGTGTGTGCACGGAGGCTGGGGAGT 435

Db 275 TCAGCCTAATCACAGATGGCTGTGAATGCTGTAAAGATATGTGCCAGAGCTTGGGGACA 334

QY 436 CCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGCCTGGTTTGTACGCTGGGGCAG 495

Db 335 ACTGCACAGAGGCTGCCATCTGTGACCCACACCGGGCCCTCTACTGCGATTACAGTGGG 394

QY 496 GGCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGA 555

Db 395 ATCGCCCGAGGTAGCAATAGGAGTGTGTGCACAGGTGGTGGTGTGGCTGTGTCTCTGG 454

QY 556 ATGGCCCGCAGGTAGCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTGTGCCGCT 615

Db 455 ATGGCGTACGCTACACCAATGGCGAGTCTCTTCCAACTGCACTGCACTGTACCT 514

QY 616 GTGATGACGG 625

Db 515 GCATTGATGG 524

RESULT 12

```
US-09-182-145-10/c
; Sequence 10, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 10
; LENGTH: 1766
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1757
; OTHER INFORMATION: Unknown base.
US-09-182-145-10

Query Match      12.3%; Score 78; DB 4; Length 1766;
Best Local Similarity 53.2%; Pred. No. 1.7e-12;
Matches 165; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 316 CCCAGCTGTGCCGGACACCCCTGTACTGTCTTGGACACCCACCCAGTGTCCACAGGGGG 375
Db 1552 CCGAATCTGCAAGTGGCCATGTGAGTGTGACCAATCCCCACCTCGCTGCCCACTGGGCG 1493

QY 376 TACCCCTGGTGGTGGATGGCTGTGGCTGTAAAGTGTGTGCACGGAGGCTGGGGGAGT 435
Db 1492 TCAGCCCTAATCACAGATGGCTGTGAATGCTGTAAAGATATGTGCCAGCAGCTTGGGGACA 1433

QY 436 CCTGGACCACTGCATGTCTGGACCCCGAGCCAGGGCTGTGTGCAGCCTGGGGCAG 495
Db 1432 ACTGCACAGAGGCTGCCATCTGTGACCCACACCGGGGCCCTCTACTGCGATTACAGTGGG 1373

QY 496 GCCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGA 555
Db 1372 ATCGCCGAGGTACGCAATAGGAGTGTGTGCACAGGTGGTGGCTGTGTCTCTGG 1313

QY 556 ATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCCTGTGCCGCT 615
Db 1312 ATGGCGTACGCTACACCAATGGCGAGTCCTTCCAACCCCACTGCAGGTACAACCTGTACCT 1253

QY 616 GTGATGACGG 625
Db 1252 GCATTGATGG 1243

RESULT 13
US-09-182-145-1
; Sequence 1, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
```

```
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 1
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-1

Query Match      11.8%; Score 75.2; DB 4; Length 2830;
Best Local Similarity 52.6%; Pred. No. 1.2e-11;
Matches 164; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 316 CCCAGCTGTGCCGGACACCCCTGTACTGTCTTGGACACCCACCCAGTGTCCACAGGGGG 375
Db 225 CCCAATCTGCAAGTGGCCATGTGAGTGTCCCGCATCCCGCCAGTGTCCCGCTGGGGG 284

QY 376 TACCCCTGGTGGTGGATGGCTGTGGCTGTAAAGTGTGTGCACGGAGGCTGGGGGAGT 435
Db 285 TCAGCCTCATCACAGATGGCTGTGAGTGTGTAAGATGTGCGCTCAGCAGCTTGGGGACA 344

QY 436 CCTGCGACCACTGCATGTCTGCGACCCCGAGCCAGGGCCCTGTTGTTCAGCCTGGGGCAG 495
Db 345 ACTGCGCGGAGGCTGCCATCTGTGACCCCGCCAGCCCGGGCCCTCTACTGTGACTACAGCGGG 404

QY 496 GCCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGA 555
Db 405 ACCGCCGAGGTACGCAATAGGAGTGTGTGCACAGGTGGTGGTGTGGGCTGCGTCCCTGG 464

QY 556 ATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCCTGTGCCGCT 615
Db 465 ATGGGCTGCGCTACAACAACGGCCAGTCCTTCCAGCCTAAGTGAAGTAACTGCACTGCACGT 524

QY 616 GTGATGACGGTG 627
Db 525 GCATCGACGGCG 536

RESULT 14
US-09-182-145-2/c
; Sequence 2, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:36:41 ; Search time 156.533 Seconds
(without alignments)
8368.853 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_635
Perfect score: 635
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues 2879534
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	635	100.0	1708	15	US-10-010-408-1 Sequence 1, Appli
2	521.4	82.1	1734	15	US-10-112-267-17 Sequence 17, Appl
3	521.4	82.1	1734	15	US-10-112-267-18 Sequence 18, Appl
4	387	60.9	753	15	US-10-010-408-3 Sequence 3, Appli
5	318	50.1	681	15	US-10-010-408-12 Sequence 12, Appl
6	278.4	43.8	1337	11	US-09-915-582-30 Sequence 30, Appl
7	278.2	43.8	1352	11	US-09-915-582-14 Sequence 14, Appl
8	273	43.0	1266	14	US-10-137-866-319 Sequence 319, App
9	273	43.0	1266	14	US-10-146-726-319 Sequence 319, App
10	273	43.0	1266	14	US-10-146-727-319 Sequence 319, App
11	273	43.0	1266	14	US-10-146-788-319 Sequence 319, App
12	273	43.0	1266	14	US-10-152-380-319 Sequence 319, App
13	273	43.0	1266	14	US-10-153-934-319 Sequence 319, App
14	273	43.0	1266	15	US-10-028-072-319 Sequence 319, App
15	273	43.0	1266	15	US-10-121-049-319 Sequence 319, App
16	273	43.0	1266	15	US-10-123-904-319 Sequence 319, App

17	273	43.0	1266	15	US-10-140-470-319	Sequence 319, App
18	273	43.0	1266	15	US-10-175-746-319	Sequence 319, App
19	273	43.0	1266	15	US-10-176-918-319	Sequence 319, App
20	273	43.0	1266	15	US-10-176-921-319	Sequence 319, App
21	273	43.0	1266	15	US-10-137-865-319	Sequence 319, App
22	273	43.0	1266	15	US-10-140-474-319	Sequence 319, App
23	273	43.0	1266	15	US-10-142-431-319	Sequence 319, App
24	273	43.0	1266	15	US-10-143-114-319	Sequence 319, App
25	273	43.0	1266	15	US-10-140-002-319	Sequence 319, App
26	273	43.0	1266	15	US-10-142-419-319	Sequence 319, App
27	273	43.0	1266	15	US-10-123-262-319	Sequence 319, App
28	273	43.0	1266	15	US-10-142-423-319	Sequence 319, App
29	273	43.0	1266	15	US-10-121-050-319	Sequence 319, App
30	273	43.0	1266	15	US-10-141-755-319	Sequence 319, App
31	273	43.0	1266	15	US-10-143-032-319	Sequence 319, App
32	273	43.0	1266	15	US-10-123-108-319	Sequence 319, App
33	273	43.0	1266	15	US-10-123-236-319	Sequence 319, App
34	273	43.0	1266	15	US-10-123-261-319	Sequence 319, App
35	273	43.0	1266	15	US-10-140-921-319	Sequence 319, App
36	273	43.0	1266	15	US-10-140-928-319	Sequence 319, App
37	273	43.0	1266	15	US-10-121-045-319	Sequence 319, App
38	273	43.0	1266	15	US-10-123-292-319	Sequence 319, App
39	273	43.0	1266	15	US-10-123-903-319	Sequence 319, App
40	273	43.0	1266	15	US-10-124-819-319	Sequence 319, App
41	273	43.0	1266	15	US-10-124-822-319	Sequence 319, App
42	273	43.0	1266	15	US-10-140-925-319	Sequence 319, App
43	273	43.0	1266	15	US-10-160-498-319	Sequence 319, App
44	273	43.0	1266	15	US-10-124-824-319	Sequence 319, App
45	273	43.0	1266	15	US-10-127-825A-319	Sequence 319, App

ALIGNMENTS

RESULT 1

US-10-010-408-1
; Sequence 1, Application US/10010408
; Publication No. US20020165185A1
; GENERAL INFORMATION:
; APPLICANT: John J. Castellet, Jr.
; TITLE OF INVENTION: NO. US20020165185A1 Heparin-Induced CCN-Like Molecules and Uses Therefor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/010,408
; FILING DATE: 07-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/044,273
; FILING DATE: March 19, 1998
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MBI-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1708 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 249..1001
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-010-408-1

Query Match 100.0%; Score 635; DB 15; Length 1708;
Best Local Similarity 100.0%; Pred. No. 4.8e-173;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCC 60
|||||
Db 1 GACGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCC 60

QY 61 GCTGGCAGTGGCTTGAATGGAGGTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 120
|||||
Db 61 GCTGGCAGTGGCTTGAATGGAGGTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 120

QY 121 CTGTCAGCTTGCTCTAAAGTCTTAGCACCTGTGGTGGCTTGGGCTTCACACACTGTGAGA 180
|||||
Db 121 CTGTCAGCTTGCTCTAAAGTCTTAGCACCTGTGGTGGCTTGGGCTTCACACACTGTGAGA 180

QY 181 CACCTTCGTGGTGGCCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGGAC 240
|||||
Db 181 CACCTTCGTGGTGGCCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGGAC 240

QY 241 ACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTCCTCTGCCTTC 300
|||||
Db 241 ACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTCCTCTGCCTTC 300

QY 301 TCTCAATGGTGTGTCCTCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCAACCC 360
|||||
Db 301 TCTCAATGGTGTGTCCTCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCAACCC 360

QY 361 AGTGCCACAGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGCTGTAAGTGTGTGCAC 420
|||||
Db 361 AGTGCCACAGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGCTGTAAGTGTGTGCAC 420

QY 421 GGAGGCTGGGGAGTCTCGACACCACTGCATGTCTCGACCCCAAGCCAGGGCCCTGGTTT 480
|||||
Db 421 GGAGGCTGGGGAGTCTCGACACCACTGCATGTCTCGACCCCAAGCCAGGGCCCTGGTTT 480

QY 481 GTCAGCCTGGGCAGGCCCTTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACG 540
|||||
Db 481 GTCAGCCTGGGCAGGCCCTTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACG 540

QY 541 GTAGCTGTGAGTGAATGGCCGACAGGTACCTGGATGGAGAGACCTTTAAACCCAATTGCA 600
|||||
Db 541 GTAGCTGTGAGTGAATGGCCGACAGGTACCTGGATGGAGAGACCTTTAAACCCAATTGCA 600

QY 601 GGGTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635
|||||
Db 601 GGGTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635

RESULT 2
US-10-112-267-17
; Sequence 17, Application US/10112267
; Publication No. US2003006878A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-112-267-17

Query Match 82.1%; Score 521.4; DB 15; Length 1734;
Best Local Similarity 92.0%; Pred. No. 2.3e-140;
Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;

QY 3 CGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCCGC 62
|||||
Db 13 CGCTCCTGATCTCCAGAGGACCCTGGGTGGGACAGGGCCCTTGGCGAGGCTGCAGCTGC 72

QY 63 TG-GGCAGTGGCTTGAATGGAGGTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCC 121
|||||
Db 73 TGTGGCAGTAGCTTGGGATGGAGGTCTTTCTGTCTGGGAACCTGAGGAGCTGAGAGGCTCC 132

QY 122 TGTGAG---CTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCA 178
|||||
Db 133 TGTGAGGCTCCTGTCTCTAAACTCTTGGCACTTGGGTGGCTTGGGCTTCACACACTGTCA 192

QY 179 GACACCTTCGTGGTGGCCCTCCACGGCCCTACCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238
|||||
Db 193 GACACCTTCGTGGTGGCCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 246

QY 239 ACACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTCTGCT 298
|||||
Db 247 ACACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCATTTCTCTCTGCT 306

QY 299 TCTCTCAATGCTGTGTCGCCAGCTGTGCCGACACCCCTGTACCTGTCTCTTGGACACCACC 358
|||||
Db 307 TCTCTCAATGCTGTATTCACAGCTGTGCCGACACCCCTGTGCTCTCTTGGACACCACC 366

QY 359 CCAGTGCCCAACAGGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTC 418
|||||
Db 367 CCAGTGCCCAACAGGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGCTGTGAGTGTGTC 426

QY 419 ACGGAGGCTGGGGAGTCTCTGCGACCACTGTCATGTCTGCGACCCCAAGGAGGCTGGT 478
|||||
Db 427 ACGGAGGCTGGGGAGTCTCTGCGACCACTGTCATGTCTGCGACCCCAAGGAGGCTGGT 486

QY 479 TTGTACGCTGGGGCAGGCCCTTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGA 538
|||||
Db 487 TTGTACGCTGGGGCAGGCCCAAGTGGCGGCTGTGTGTGCTCTCTGAAAGAGGATGA 546

QY 539 CGGTAGCTGTGAGGTGAATGGCCGACAGGTACCTGGATGGAGAGACCTTTAAACCCAATTG 598
|||||
Db 547 CGGGAGCTGTGAGGTGAATGGCCGACAGGTACCTGGATGGGAGAGACCTTTAAACCCAATTG 606

QY 599 CAGGGTCTCTGCGGCTGTGATGACGGTGGCTTCACC 635
|||||
Db 607 CAGGGTTTGTGCCGCTGTGATGACGGTGGCTTCACC 643

RESULT 3
US-10-112-267-18/c
; Sequence 18, Application US/10112267

Publication No. US20030068678A1
GENERAL INFORMATION:
APPLICANT: Botstein, David A.
APPLICANT: Cohen, Robert
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Lawrence, David A.
APPLICANT: Levine, Arnold J.
APPLICANT: Pennica, Diane
APPLICANT: Roy, Margaret Ann
APPLICANT: Wood, William I.
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: P1176R2
CURRENT APPLICATION NUMBER: US/10/112,267
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
NUMBER OF SEQ ID NOS: 156
SEQ ID NO 18
LENGTH: 1734
TYPE: DNA
ORGANISM: Mus musculus
US-10-112-267-18

Query Match 82.1%; Score 521.4; DB 15; Length 1734;
Best Local Similarity 92.0%; Pred. No. 2.3e-140;
Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;

QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGC 62
Db 1722 CGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCCCTTGGCGAGGCTGCAGCTGC 1663

QY 63 TG-GGCAGTGGCTTGGAAATGGAGGTCTTTATTACTGGGAACTGAGGAGCTAAGAGGCTCC 121
Db 1662 TGTGGCAGTAGCTTGGGATGGAGGTCTTTCTTGTGGGAACTGAGGAGCTGAGAGGCTCC 1603

QY 122 TGTGAG--CTTGTCTCTAAAGTCTTAGCACCTTGTGGTGGCTTGGGCTTCACACACTGTCA 178
Db 1602 TGTGAGGCTCCTGTCTCTAAACTCTTGGCACCTTGGGTGGCTTGGGCTTCACACACTGTCA 1543

QY 179 GACACCTTCGTGGTGGCTCCACGGCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238
Db 1542 GACACCTTCTTGGTGGCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 1489

QY 239 ACACGGTGACATGAGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCCTTCCTCTGCCT 298
Db 1488 ACACGGTGACATGAGGGCAACCCCACTGATCCATCTTCTGGCCATTTCCCTTCCTCTGCAT 1429

QY 299 TCTCTCAATGGTGTGCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCACC 358
Db 1428 TCTCTCAATGGTGTATCCAGCTGTGCCCAGCACCCCTGTGCCTGTCCCTTGGACACCACC 1369

QY 359 CCAGTGCCCAACAGGGGTACCCCTGGTGGCTGGATGGCTGTGGCTGTCTGTAAGTGTGTGC 418
Db 1368 CCAGTGCCCAACCGGGGTACCCCTGGTGGATGGCTGTGGCTGTCTGTCGAGTGTGTGC 1309

QY 419 ACGGAGGCTGGGGAGTCCTGCGACCACTGCATGTCTGCGACCCCAAGCCAGGCGCCTGGT 478
Db 1308 ACGGAGGCTGGGGAGTCCTGCGACCACTGCATGTCTGCGACCCCAAGCCAGGCGCCTGGT 1249

QY 479 TTGTACGCTGGGGCAGGCCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGA 538
Db 1248 TTGTACGCTGGGGCAGGCCCAGTGGCCGCTGGTGTGTGTCTCTTCGAAGAGGATGA 1189

QY 539 CGGTAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCAATTG 598
Db 1189 CGGTAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCAATTG 598

Db 1188 CGGAGCTGTGAGGTGAATGGCCGCGAGGTACTGGATGGGAGACCTTTAAACCAATTG 1129

QY 599 CAGGGTCTGTGCCGCTGTGATGACGGTGGCTTCACC 635
Db 1128 CAGGGTTTGTGCCGCTGTGATGACGGTGGTTTCACC 1092

RESULT 4
US-10-010-408-3
Sequence 3, Application US/10010408
Publication No. US20020165185A1
GENERAL INFORMATION:
APPLICANT: John J. Castellet, Jr.
TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CCN-Like Molecules
and Uses Therefor
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 753 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..750
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-010-408-3

Query Match 60.9%; Score 387; DB 15; Length 753;
Best Local Similarity 100.0%; Pred. No. 1e-101;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ATGAGGGGCAGCCCACTGATCCATCTTCTGGCACCTTCTCCTCTGCTCTCTCAATG 308
Db 1 ATGAGGGGCAGCCCACTGATCCATCTTCTGGCACCTTCTCCTCTGCTCTCAATG 60

QY 309 GTGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCCCA 368
Db 61 GTGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCCCA 120

QY 369 CAGGGGGTACCCCTGCTGGATGGCTGTGGCTGTCTGTAAGTGTGTGACGAGGCTG 428
Db 121 CAGGGGGTACCCCTGCTGGATGGCTGTGGCTGTCTGTAAGTGTGTGACGAGGCTG 180

Db 64 TCAAAGGTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCCGA 123
QY 363 TGCCACACAGGGGTACCCCTGGTGGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTCACGG 422
Db 124 TGCCCGCTGGAGTACCCCTGGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG 183
QY 423 AGGCTGGGGAGTCCCTGCGACCACTTCACATGTCTGCGACCCAGCCAGGGCCTGGTTGT 482
Db 184 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGC 243
QY 483 CAGCCTGGGGAGCCCTGCGACCACTTCACATGTCTGCGACCCAGCCAGGGCCTGGTTGT 542
Db 244 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGC 303
QY 543 AGCTGTGAGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363
QY 603 GTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635
Db 364 ATCCGCTGCGCTGCGAGGACGGCGGCTTCACC 396

RESULT 9

US-10-146-726-319
; Sequence 319, Application US/10146726
; Publication No. US20030129690A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C308
; CURRENT APPLICATION NUMBER: US/10/146,726
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-726-319

Query Match 43.0%; Score 273; DB 14; Length 1266;
Best Local Similarity 80.9%; Pred. No. 6.8e-69;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGGAGCCACTGTGCCGACACCCCTGTACCTGTCTGTCCTCTGCTCTC 302
Db 4 GGGACATGAGAGGCACACCGAAGACCCACTCTGGCCTTCTCCCTCTGCTCTCCTC 63
QY 303 TCAATGGTGTGTCGCCAGCTGTGCCGACACCCCTGTACCTGTCTGGACACCCACCCAG 362
Db 64 TCAAAGGTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCCGA 123
QY 363 TGCCACACAGGGGTACCCCTGGTGGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTCACGG 422

Db 124 TGCCCGCTGGGAGTACCCCTGGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG 183
QY 423 AGGCTGGGGAGTCCCTGCGACCACTTCACATGTCTGCGACCCAGCCAGGGCCTGGTTGT 482
Db 184 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGC 243
QY 483 CAGCCTGGGGAGCCCTGCGGCACTTCACATGGGGCTGTGTGCTCTTGGATGAGGATGACGGT 542
Db 244 CAGCCCGGGCAGGACCCCGTGGCGGGGGCCCTGTGCCCTCTTGGCAGAGGACGACAGC 303
QY 543 AGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363
QY 603 GTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635
Db 364 ATCCGCTGCGCTGCGAGGACGGCGGCTTCACC 396

RESULT 10

US-10-146-727-319
; Sequence 319, Application US/10146727
; Publication No. US20030129691A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C312
; CURRENT APPLICATION NUMBER: US/10/146,727
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-727-319

Query Match 43.0%; Score 273; DB 14; Length 1266;
Best Local Similarity 80.9%; Pred. No. 6.8e-69;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGGAGCCACTGTGCCGACACCCCTGTACCTGTCTGTCCTCTGCTCTC 302
Db 4 GGGACATGAGAGGCACACCGAAGACCCACTCTGGCCTTCTCCCTCTGCTCTCCTC 63
QY 303 TCAATGGTGTGTCGCCAGCTGTGCCGACACCCCTGTACCTGTCTGGACACCCACCCAG 362
Db 64 TCAAAGGTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCCGA 123
QY 363 TGCCACACAGGGGTACCCCTGGTGGATGGCTGTGGCTGCTGTAAAGTGTGTCACGG 422
Db 124 TGCCCGCTGGGAGTACCCCTGGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG 183
QY 423 AGGCTGGGGAGTCCCTGCGACCACTTCACATGTCTGCGACCCAGCCAGGGCCTGGTTGT 482
Db 184 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGC 243

QY 483 CAGCCTGGGGCAGGCCCTGGCGGCCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542
Db 244 CAGCCCGGGCAGACCCGGTGGCGGGCCCTGTGCCTCTTGGCAGAGGACGACAGC 303
QY 543 AGCTGTGAGGTGAATGGCCGACAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363
QY 603 GTCCTGTGCCGCTGTGTGATGACGGTGGCTTCACC 635
Db 364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACC 396

RESULT 11
US-10-146-788-319
; Sequence 319, Application US/10146788
; Publication No. US20030129693A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC322
; CURRENT APPLICATION NUMBER: US/10/146,788
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-788-319

Query Match 43.0%; Score 273; DB 14; Length 1266;
Best Local Similarity 80.9%; Pred. No. 6.8e-69;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTCCTCTGCCTTCTC 302
Db 4 GGGGACATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCTCTCTCTCTCTC 63
QY 303 TCAATGGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTGTAAGTGTGTGCACGG 362
Db 64 TCAAAGGTGCTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGCCACCTCCCCGA 123
QY 363 TGCCACAGGGGTACCCCTGGTGGTGGCTGTGGCTGTGGCTGTGTAAGTGTGTGCACGG 422
Db 124 TGCCCGCTGGAGTACCCCTGGTGGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 183
QY 423 AGGCTGGGGAGTCCCTGCGACACCTGCATGTCTGCGACCCACGACCCAGGCGCTTGT 482
Db 184 CGGCTGGGGAGTCCCTGCGACCAACTCCACGTCTCGACGCCAGCCAGGCTTGTCTGC 243
QY 423 AGGCTGGGGAGTCCCTGCGACACCTGCATGTCTGCGACCCACGACCCAGGCGCTTGT 482
Db 184 CGGCTGGGGAGTCCCTGCGACCAACTCCACGTCTCGACGCCAGCCAGGCGCTTGTCTGC 243
QY 423 AGGCTGGGGAGTCCCTGCGACACCTGCATGTCTGCGACCCACGACCCAGGCGCTTGT 482
Db 184 CGGCTGGGGAGTCCCTGCGACCAACTCCACGTCTCGACGCCAGCCAGGCGCTTGTCTGC 243
QY 483 CAGCCTGGGAGCCCTTCCGACCACTCCACGTCTGCGACGCCAGCCAGGCGCTTGTGT 542
Db 244 CAGCCCGGGCAGGACCCGGTGGCCGGCCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 303

QY 543 AGCTGTGAGGTGAATGGCCGACAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363
QY 603 GTCCTGTGCCGCTGTGTGATGACGGTGGCTTCACC 635
Db 364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACC 396
RESULT 12
US-10-152-380-319
; Sequence 319, Application US/10152380
; Publication No. US20030129694A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC397
; CURRENT APPLICATION NUMBER: US/10/152,380
; CURRENT FILING DATE: 2002-05-21
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-152-380-319

Query Match 43.0%; Score 273; DB 14; Length 1266;
Best Local Similarity 80.9%; Pred. No. 6.8e-69;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTCCTCTGCCTTCTC 302
Db 4 GGGGACATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCTCTCTCTCTCTC 63
QY 303 TCAATGGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTGTAAGTGTGTGCACGG 362
Db 64 TCAAAGGTGCTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGCCACCTCCCCGA 123
QY 363 TGCCACAGGGGTACCCCTGGTGGTGGCTGTGGCTGTGGCTGTGTAAGTGTGTGCACGG 422
Db 124 TGCCCGCTGGAGTACCCCTGGTGGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 183
QY 423 AGGCTGGGGAGTCCCTGCGACACCTGCATGTCTGCGACCCACGACCCAGGCGCTTGT 482
Db 184 CGGCTGGGGAGTCCCTGCGACCAACTCCACGTCTCGACGCCAGCCAGGCTTGTCTGC 243
QY 483 CAGCCTGGGAGCCCTTCCGACCACTCCACGTCTGCGACGCCAGCCAGGCGCTTGTGT 542
Db 244 CAGCCCGGGCAGGACCCGGTGGCCGGCCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 303
QY 543 AGCTGTGAGGTGAATGGCCGACAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363
QY 603 GTCCTGTGCCGCTGTGTGATGACGGTGGCTTCACC 635

Db 364 ATCCGCTGCGCTGCGAGGACGGCGGCTTCACC 396

RESULT 13

US-10-153-934-319

; Sequence 319, Application US/10153934

; Publication No. US20030129695A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C412

; CURRENT APPLICATION NUMBER: US/10/153,934

; CURRENT FILING DATE: 2002-05-22

; Prior Application removed - See file Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 319

; LENGTH: 1266

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-153-934-319

Query Match 43.0%; Score 273; DB 14; Length 1266;

Best Local Similarity 80.9%; Pred. No. 6.8e-69;

Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACCTTCTTCTCTGCTGCTCTC 302

Db 4 GGGACATGAGAGGCACACCGAAGACCCACCTCTCTGGCCCTTCTCCCTCCTCTGCTCTC 63

QY 303 TCAATGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCAACCCACG 362

Db 64 TCAAAGTGGGTACCCAGCTGTGCCGACACCATGTACCTGCTGCTGCTGCTGCTGCTGCTG 123

QY 363 TGCCACAGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTGTAAGTGTGTGCACGG 422

Db 124 TGCCCGTGGGAGTACCCCTGGTGGTGGATGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTG 183

QY 423 AGGCTGGGGAGTCTCGACCACTGCATGTCTGCGACCCAGCCAGGCTGTTGT 482

Db 184 CGGCTGGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGGCGCTGGTCTGC 243

QY 483 CAGCCTGGGGCAGGCCCTGGCGCCATGGGCTGTGTGTCTTGGATGAGGATGACGGT 542

Db 244 CAGCCCGGGCAGGACCCGGTGGCGGGGGCCCTGTGCTCTTGGCAGGACGACAGC 303

QY 543 AGCTGTGAGGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602

Db 304 AGCTGTGAGGTGAACGGCCCGCTGTATCGGAAGGGAGACCTTCCAGCCCCACTGCAGC 363

QY 603 GTCCTGTCCCGTGTGATGACGGTGGCTTCACC 635

Db 364 ATCCGCTGCCGTGCGAGGACGGCGGCTTCACC 396

RESULT 14

US-10-028-072-319
; Sequence 319, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29

Db 184 CCGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGGCGCCTGGTCTGC 243
QY 483 CAGCCTGGGGCAGGCCCTGGCGGCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT 542
Db 244 CAGCCCGGGCAGGACCCGGTGGCCGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGC 303
QY 543 AGCTGTGAGGTGAATGGCCGAGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGGTGAACGGCCGCCCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363
QY 603 GTCCTGTGCCGCTGTGTATGATGACGGTGGCTTCACC 635
Db 364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACC 396

RESULT 15

US-10-121-049-319
; Sequence 319, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-121-049-319

Query Match 43.0%; Score 273; DB 15; Length 1266;
Best Local Similarity 80.9%; Pred. No. 6.8e-69;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCTTCCCTGCTTC 302
Db 4 GGGACATGAGAGGCACACCGAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCTCCTC 63
QY 303 TCAATGCTGTGTGCCAGCTGTGCCGGACACCCTGTACCTGTCTTGGACACCCACCCAG 362
Db 64 TCAAAGGTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCCGA 123
QY 363 TGCCACAGGGGGTACCCCTGGTCTGGATGGCTGTGGCTGCTGTAAAGTGTGTCACGG 422
Db 124 TGCCCGCTGGGAGTACCCCTGGTGGATGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 183
QY 423 AGGTGGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGCGCTGTTGT 482
Db 184 CGGCTGGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGACCCAGGCGCTGGTCTGC 243
QY 483 CAGCCTGGGGCAGGCCCTGGCGGCATGGGGCTGTGTCTCTTGGATGAGGATGACGCT 542

Db 244 CAGCCCGGGGCAGGACCCCGTGGCCGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGC 303
QY 543 AGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGGTGAACGGCCGCCCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363
QY 603 GTCCTGTGCCGCTGTGTATGATGACGGTGGCTTCACC 635
Db 364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACC 396

Search completed: July 28, 2003, 21:31:54
Job time : 157.533 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 18:02:17 ; Search time 1126 Seconds
(without alignments)
9133.322 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_635
Perfect score: 635
Sequence: 1 GACGGCTTCTGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_mam:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	425.8	67.1	940	BQ937887	BQ937887 AGENCOURT
2	385.2	60.7	537	BQ560868	BQ560868 H4067A01-
3	335.8	52.9	424	BB849097	BB849097 BB849097
4	299	47.1	749	AL555144	AL555144 AL555144
5	297.4	46.8	1166	BM543799	BM543799 AGENCOURT
6	271.4	42.7	1251	BQ961357	BQ961357 AGENCOURT

7	269.4	42.4	452	10	BE481184	BE481184
8	267.6	42.1	1058	14	BM805088	BM805088 AGENCOURT
9	267.2	42.1	620	13	BG928868	BG928868 HNC57-1-D
10	262	41.3	750	13	BM043988	BM043988 603620978
11	261.6	41.2	1006	14	BM921531	BM921531 AGENCOURT
12	258.2	40.7	928	13	BI161474	BI161474 602864871
13	258.2	40.7	979	14	BQ279131	BQ279131 AGENCOURT
14	256.8	40.4	933	14	BQ278961	BQ278961 AGENCOURT
15	256	40.3	1073	14	BQ073722	BQ073722 AGENCOURT
16	247.8	39.0	800	13	BI826781	BI826781 603077268
17	246.2	38.8	888	13	BI822562	BI822562 603072631
18	245	38.6	886	13	BI822142	BI822142 603039845
19	238.2	37.5	651	13	BI457141	BI457141 603185392
20	229	36.1	916	13	BI457367	BI457367 603185689
21	223.2	35.1	405	12	BG900069	BG900069 HOA51-1-A
22	209.6	33.0	380	12	BG900020	BG900020 HOA48-1-G
23	205.8	32.4	618	12	BG538695	BG538695 602566932
24	205.2	32.3	792	13	BI823598	BI823598 603040962
25	161.2	25.4	657	13	BM488499	BM488499 pgm2n.pk0
26	145.4	22.9	790	13	BM046275	BM046275 603626068
27	144.4	22.7	1022	14	BQ952960	BQ952960 AGENCOURT
28	117.6	18.5	742	13	BI758148	BI758148 603023866
29	110.2	17.4	190	9	AA647775	AA647775 vp04c09.r
30	92.8	14.6	502	10	AV670831	AV670831 AV670831
31	87.2	13.7	828	9	AL543019	AL543019 AL543019
32	85.6	13.5	505	12	BF037009	BF037009 601456766
33	85.6	13.5	622	14	BM772255	BM772255 K-EST0056
34	85.6	13.5	756	9	AU125723	AU125723 AU125723
35	85.6	13.5	826	12	BG116760	BG116760 602318826
36	85.6	13.5	827	9	AU131392	AU131392 AU131392
37	85.6	13.5	851	13	BI828218	BI828218 603073692
38	85.6	13.5	874	14	BQ720581	BQ720581 AGENCOURT
39	85.6	13.5	877	9	AL546864	AL546864 AL546864
40	85.6	13.5	878	14	BQ719477	BQ719477 AGENCOURT
41	85.6	13.5	887	14	BQ230898	BQ230898 AGENCOURT
42	85.6	13.5	888	9	AL544380	AL544380 AL544380
43	85.6	13.5	896	14	BQ719732	BQ719732 AGENCOURT
44	85.6	13.5	931	14	BQ880701	BQ880701 AGENCOURT
45	85.6	13.5	942	9	AL544453	AL544453 AL544453

ALIGNMENTS

RESULT 1
BQ937887
LOCUS BQ937887 940 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8951807 NCI_CGAP_Co24 Mus musculus cDNA clone
IMAGE:6476852 5', mRNA sequence.
ACCESSION BQ937887
VERSION BQ937887.1 GI:22353365
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM14017 row: n column: 21
High quality sequence stop: 543.
FEATURES Location/Qualifiers
source 1. .940

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6476852"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: Sali; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      169 a   277 c   288 g   200 t      6 others
ORIGIN

Query Match      67.1%; Score 425.8; DB 14; Length 940;
Best Local Similarity 90.9%; Pred. No. 1.5e-99;
Matches 488; Conservative 0; Mismatches 39; Indels 10; Gaps 3;

QY 1 GACGCTCTGATCTCCAGAGGACCTGGGGTGGGACAGGGGCTGGCAAGGCTGCAGCC 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37 GACGCTCTGATCTCCAGAGGACCCGGGCTGGGACAGGGGCTGGCGAGGCTGCAGCT 96

QY 61 GCTG-GGCAGTGGCTTGGAAATGGAGTCTTTATTACTGGGAAGTGAAGAGGCT 119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 GCTGTGGCAGTAGCTTGGGATGGAGTCTTTCTTGTGGGAAGTGAAGAGGCT 156

QY 120 CCTGTGAG--CTTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGT 176
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 CCTGTGAGGCTCTCTGCTCTAAACTCTTGGCACTTGGGCTGGGCTTCACACACTGT 216

QY 177 CAGACACCTTCGTGGTGGCTCCACGGGCTCACCTTCAGGTTGAAGCTGGCTCCACAAG 236
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 CAGACACCTTCTTGGTGGCTCTCTGGCC-----TCAGGTTGAAGCTGGCTCCACAAG 270

QY 237 GGACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGTC 296
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 GGACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGTC 330

QY 297 CTTCTCTCAATGGTGTGCCCCAGCTGTGCCGAGACACCCCTGACCTGTCTGTAAGTGTGT 356
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 ATTCTCTCAATGGTGTATGCCAGCTGTGCCAGCACCCCTGTGGCTGTCTGTAAGTGTGT 450

QY 357 CCCCAGTGGCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAGTGTGT 416
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 CCCCAGTGGCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAGTGTGT 450

QY 417 GCACGGAGCTGGGGGAGTCTCGACACCACTGCATGTCTGCCACCCAGCCAGGGGCTG 476
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 GCACGGAGCTGGGGGAGTCTCGACACCACTGCATGTCTGCCACCCAGCCAGGGGCTG 510

QY 477 GTTTGTGACCTGGGGCAGGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAG 533
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 GTTTGTGACCTGGGGCAGGGCCCTGGCGGGTGGTGTGTGTGTGTGTGTGTGTGTGTGT 567

RESULT 2
BQ560868
LOCUS      BQ560868      537 bp      mRNA      linear      EST 20-JUN-2002
DEFINITION H4067A01-5 NIA Mouse 7.4K cDNA clone Set Mus musculus cDNA clone
            H4067A01 5', mRNA sequence.
ACCESSION  BQ560868
VERSION    BQ560868.1 GI:21461753
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 537)
AUTHORS   VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G., Martin
            ,P.R., Stagg,C.A., Bassey,U., Aiba,K., Hamatani,T., Kargul,G.J.,
            Luo,A.G. and Ko,M.S.H.
TITLE     Assembly, verification, and initial annotation of NIA 7.4K mouse
            cDNA clone set
JOURNAL   Unpublished (2002)

```

```

COMMENT      Contact: Yong Qian
              Laboratory of Genetics
              National Institute on Aging/National Institutes of Health
              333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
              Email: cdna@igsun.grc.nia.nih.gov
              This clone set has been freely distributed to the community. Please
              visit http://igsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html for details.
              Plate: H4067 Row: A Column: 01
              Seq primer: -21M13 Reverse
              High quality sequence stop: 537
              POLYA=No.

FEATURES             Location/Qualifiers
     source            1..537
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="niaEST:H4067A01-5"
                     /db_xref="taxon:10090"
                     /clone="H4067A01"
                     /clone_lib="NIA Mouse 7.4K cDNA Clone Set"
                     /sex="mixed"
                     /dev_stage="mixed"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Sali; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."

BASE COUNT      87 a   162 c   166 g   122 t
ORIGIN

Query Match      60.7%; Score 385.2; DB 14; Length 537;
Best Local Similarity 92.1%; Pred. No. 3.7e-89;
Matches 442; Conservative 0; Mismatches 28; Indels 10; Gaps 3;

QY 1 GACGCTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCC 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 GACGCTCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCTTGGCGAGGCTGCAGCT 123

QY 61 GCTG-GGCAGTGGCTTGGAAATGGAGTCTTTATTACTGGGAAGTGAAGAGGCT 119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 GCTGTGGCAGTAGCTTGGGATGGAGTCTTTCTTGTGGGAAGTGAAGAGCT 183

QY 120 CCTGTGAG--CTTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGT 176
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 CCTGTGAGGCTCTGCTCTCTAAACTCTTGGCACTTGGGCTGGGCTTCACACACTGT 243

QY 177 CAGACACCTTCGTGGTGGCTCCACGGGCTCACCTTCAGGTTGAAGCTGGCTCCACAAG 236
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 CAGACACCTTCTTGGTGGCTCTCTCGGCC-----TCAGGTTGAAGCTGGCTCCACAAG 297

QY 237 GGACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGTC 296
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 298 GGACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGTC 357

QY 297 CTTCTCTCAATGGTGTGTGCCAGCTGTGCCGAGACACCCCTGACCTGTACCTTGGACACCA 356
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 358 ATTCTCTCAATGGTGTATGCCAGCTGTGCCAGACCCCTGTGGCTGTCTTGGACACCA 417

QY 357 CCCCAGTGGCCAGAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAGTGTGT 416
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 418 CCCCAGTGGCCAGAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAGTGTGT 477

QY 417 GCACGGAGGCTGGGGGAGTCTCTGGACCACTGCATGTCTGCCACCCAGCCAGGGGCTG 476
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 478 GCACGGAGGCTGGGGGAGTCTCTGGACCACTGCATGTCTGCCACCCAGCCAGGGGCTG 537

RESULT 3
BB849097
LOCUS      BB849097      424 bp      mRNA      linear      EST 26-NOV-2001
DEFINITION BB849097 RIKEN full-length enriched, adult inner ear Mus musculus
            cDNA clone F93006G02 5', mRNA sequence.
ACCESSION  BB849097
VERSION    BB849097.1 GI:17090551
KEYWORDS   EST.

```


SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
FEATURES Location/Qualifiers
source 1. .424
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930006G02"
/clone_lib="RIKEN full-length enriched, adult inner ear"
/tissue_type="adult"
/dev_stage="adult"
/note="pooled tissues ; (tissue_type=cerebellum, dev_stage=16 days neonate, sex=mixed), (tissue_type=cerebellum, dev_stage=0 day neonate, sex=mixed), (tissue_type=hippocampus, dev_stage=adult, sex=male), (tissue_type=whole body, dev_stage=9 days embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days embryo, sex=mixed)"
BASE COUNT 65 a 124 c 132 g 103 t
ORIGIN
Query Match 52.9%; Score 335.8; DB 10; Length 424;
Best Local Similarity 91.4%; Pred. No. 2e-76;
Matches 392; Conservative 0; Mismatches 27; Indels 10; Gaps 3;
QY 37 AGGGCCCTTGGCAAGGCTGCAGCCGCTG-GGCAGTGGCTTGGATGGAGGCTCTTTATTAC 95
Db 2 AGGGCCCTTGGCAGGCTGCAGCTGGTGGCAGTAGCTTGGGATGGAGGCTCTTTCTTGC 61
QY 96 TGGGAAGTGAAGAGCTAAGAGGCTCCTGTGTCAG---CTGTCTAAAGTCTTAGCACTTGT 152

Db 62 TGGGAAGTGAAGAGCTGAGAAGCTCCTGTTCAGGCTCCTGTCTAAACTCTTGGCACTTGC 121
QY 153 GGTGGCTTGGGCTTTCACACACTGTCAGACACACTTTCGTGGTGGCTTCCACGGCCTCACCTT 212
Db 122 GGTGGCTTGGGCTTTCACACACTGTCAGACACACTTCTTGGTGGCTCCTCGGCC-----T 175
QY 213 CAGGTTGAAGCTGGCTCCACACAGGACACGGTGACATGAGGGGACGCCCACTGATCCAT 272
Db 176 CAGGTTGAAGCTGGCTCCACACAGGACACGGTGACATGAGGGGACGCCCACTGATCCAT 235
QY 273 CTTCTGGCCACTTCCCTTCCCTCTCTCTCAATGGTGTGCCCCAGCTGTGCCGGACA 332
Db 236 CTTCTGGCCACTTCCCTTCCCTCTCTCTCAATGGTGTATGCCAGCTGTGCCAGCA 295
QY 333 CCCTGTACCTGTCTTGGACACACCCCACTGCCCCACAGGGGTACCCCTGGTGGTGGAT 392
Db 296 CCCTGTGCTGTCTTGGACACACCCCACTGCCCCACCGGGGTACCCCTGGTGGTGGAT 355
QY 393 GGCTGTGGCTGTGTAAAGTGTGTGCACGGAGGCTGGGGGAGCTCTGCGACCACTGTCAT 452
Db 356 GGCTGTGGCTGTGTGCGAGTGTGTGCACGGAGGCTGGGGGAGTCTTGGACCACTGTCAT 415
QY 453 GTCTGGCAG 461
Db 416 GTCTGCAAC 424
RESULT 4
AL555144
LOCUS AL555144 LTI_NFL006_PL2 749 bp mRNA linear EST 16-FEB-2001
DEFINITION AL555144 prime, mRNA sequence.
ACCESSION AL555144
VERSION AL555144.1 GI:12896595
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 749)
JOURNAL Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DK007Y021"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 128 a 252 c 236 g 133 t
ORIGIN
Query Match 47.1%; Score 299; DB 9; Length 749;
Best Local Similarity 79.1%; Pred. No. 8.7e-67;
Matches 371; Conservative 0; Mismatches 90; Indels 8; Gaps 1;
QY 167 CACACACTGTTCAGACACCTTTCGTGGTGGCTCCACGGCCTCACCTTCAGGTTTGAAGCTG 226

||||| | ||| |||| | ||||| ||| | ||||| |||| | ||||| ||||| | ||||| 291
Db 232 CACACAGGACAGGACACCCCTTGGTGGCCTTACAGTTTACCTTCAGGCTCAAAGCTG 291

||||| ||| ||| ||||| ||||| ||| ||| ||||| ||||| ||| ||| ||||| ||||| 286
QY 227 GCTCCACAAGGACACAGGTGACATGAGGGGACGCCCACTGATCCATCTTGTGGCCACTTC 286

||||| ||| ||| ||||| ||||| ||| ||| ||||| ||||| ||| ||| ||||| ||||| 343
Db 292 GCTCTGCAGGG-----GACATGAGAGGACACCGAAGACCCACCTCCTGGCCTTCTC 343

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 346
QY 287 CTTCTCTGCCTTCTCTCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCC 346

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 403
Db 344 CCTCCTCTGCCTCTCTCAAGGTGCGTACCCAGCTGTGCCCGACACCACTGTACCTGCC 403

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 406
QY 347 TTGGACACCAACCCAGTCCCCACAGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGCTG 406

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 463
Db 404 CTGGCCACCTCCCGATGCCCGTGGGAGTACCCCTGGTGGATGGCTGTGGCTGCTG 463

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 466
QY 407 TAAAGTGTGTCACGAGGCTGGGGGAGTCTTCGACCACTGCATGCTGTGGACCCAG 466

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 523
Db 464 CCGGGTATGTGACGCGCGGCTGGGGAGCCCTGCGACCAACTCCACGCTTCCGACGCCAG 523

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 526
QY 467 CCAGGGCCTGTTGTACAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTT 526

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 583
Db 524 CCAGGGCCTGGTCTGCCAGCCCGGGCAGGACCCGGTGGCCGGGGCCTGTGCCCTCTT 583

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 586
QY 527 GGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTT 586

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 643
Db 584 GGCAGAGGACGACAGCAGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTT 643

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 635
QY 587 TAAACCCCAATTGCAGGGTCTGTGCCGCTGTGATGACGGTGGCTTTCACC 635

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 692
Db 644 CCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGGCTTTCACC 692

RESULT 5
BM543799
LOCUS
DEFINITION
5', mRNA sequence.
BM543799
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1166)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12361 row: f column: 15
High quality sequence stop: 547.
Location/Qualifiers
1. .1166
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5589134"
/clone_lib="NIH_MGC_125"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcorV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcorV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was

constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT 216 a 424 c 315 g 207 t 4 others

ORIGIN

Query Match 46.8%; Score 297.4; DB 13; Length 1166;
Best Local Similarity 78.9%; Pred. No. 2.7e-66;
Matches 370; Conservative 0; Mismatches 91; Indels 8; Gaps 1;

QY 167 CACACACTGTCAGACACACCTTCGTGGTGGCTCCACGGCCTCACCTTCAGGTTTGAAGCTG 226
||||| | ||| |||| | ||||| ||| | .||||||| | |||||
Db 137 CACACACGGACAGCACCCCTTGGTGGCCTTCACAGTTTCACTTCAGGCTCAAAGCTG 196

QY 227 GCTCCACAAGGACACGGTGACATGAGGGGACGCCCACTGATCCATCTTCTGGGCCACTTC 286
|||| | || | ||||| |||| | || | || | || | || | || | || | || | || | || |
Db 197 GCTCTGCAGGG-----GACATGAGAGGACACCCGAGACCCACTCTCGCCTTCTC 248

QY 287 CTTCTCTGCCTTCTCTCAATGGTGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTC 346
| ||||| ||||| ||||| |||| | || | ||||| ||||| ||||| ||||| ||||| |||||
Db 249 CCTCCTCTGCCTCTCTCAAAGGTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCC 308

QY 347 TTGGACACCAACCCAGTCCCCACAGGGGTACCCCTGGTGTGGATGGCTGGCTGCTG 406
|| | |||| | |||| | || | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 CTGGCCACCTCCCGATGCCCGTGGGAGTACCCCTGGTGTGGATGGCTGGCTGCTG 368

QY 407 TAAAGTGTGTCACGGAGGCTGGGGGAGTCTTCGACCACTGCATGTCTGCGACCCAG 466
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 369 CCGGGTATGTGCACGGCGGCTGGGGGAGCCCTGCGACCAACTCCACGTTGCGACGCCAG 428

QY 467 CCAGGGCCTGGTTTGTACAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTT 526
||||| ||||| || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 429 CCAGGGCCTGGTCTGCCAGCCCGGGCAGGACCCGGTGGACGGGGGCGCTTTCACC 488

QY 527 GGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTT 586
|| | |||| | || | ||||| ||||| ||||| |||| | || | || | || | || | || | || |
Db 489 GGCAGAGGACGAGAGCAGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTT 548

QY 587 TAAACCCCAATTGAGGGTCTGTGCCGCTGTGATGACGGTGGCTTTCACC 635
| || | |||| | || | ||||| || | ||||| || | ||||| ||||| ||||| ||||| |||||
Db 549 CCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGGCTTTCACC 597

RESULT 6
BQ961357
LOCUS
DEFINITION
BQ961357
1251 bp mRNA linear EST 21-AUG-2002
AGENCOURT_8929398 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6484532
5', mRNA sequence.
BQ961357
VERSION
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1251)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2670 row: n column: 21
High quality sequence stop: 465.
Location/Qualifiers
1. .1251
/organism="Homo sapiens"
/db_xref="taxon:9606"

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source


```
Db 193 CCACGTCGCGACGCCAGCCAGGCGCCTGGTCTGCCAGCCCGGGGCAGGACCGGTCGACG 252
QY 509 TGGGCTGTGTGTCTCTTTGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGCAGGTA 568
Db 253 GGGGCGCCTGTGCTCTTTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGGCCGCTGTA 312
QY 569 CCTGATGGAGAGACCTTTAAACCCAAATTGCAGGTCCTGTGTCGCCGCTGTGATGACGGTGG 628
Db 313 TCGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGG 372
QY 629 CTTACCC 635
Db 373 CTTACCC 379

RESULT 14
BQ278961
LOCUS
DEFINITION BQ278961 933 bp mRNA linear EST 07-MAY-2002
AGENCOURT_7046697 NIH_MGC_107 Homo sapiens cDNA clone IMAGE:5805626
5', mRNA sequence.
ACCESSION BQ278961
VERSION BQ278961.1 GI:20489169
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 933)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2045 row: c column: 03
High quality sequence start: 12
High quality sequence stop: 353.
FEATURES
Location/Qualifiers
source
1. .933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5805626"
/clone_lib="NIH_MGC_107"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 139 a 330 C 295 g 166 t 3 others
ORIGIN
Query Match 40.4%; Score 256.8; DB 14; Length 933;
Best Local Similarity 78.9%; Pred. No. 8e-56;
Matches 306; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 248 CATGAGGGGCGACCCACTGATCCATCTCTTGGCCACTTCCTCTCTCTCTCTCAAT 307
Db 16 CAAGACGAGGGCACACCCGAAACCCCTCTCTGGCCTTCTCCCTCTCTCTCTCTCAAA 75
QY 308 GGTGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCCC 367
Db 76 GGTGGTACCCAGCTGTGTCCGACACCATGTACCTGCCCTGGCCACCTCCCCGATGCC 135
```

```
QY 368 ACAGGGGGTACCCCTGGTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGGCT 427
Db 136 GCTGGGAGTACCCCTGGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGGCGGCT 195
QY 428 GGGGGAGTCTCTGCGACCACTGTCATGTCTGCGACCCCGAGGCGCCTGGTTGTAGCC 487
Db 196 GGGGGAGCCCTGCGACCACTCCACGTCTGCGACCGCCAGGCGCCTGGTCTGCCAGCC 255
QY 488 TGGGGCAGGCCCTGGCGCCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTG 547
Db 256 CGGGCAGGACCCGGTGGACGGGGGGCCCTGTGCCCTTGTTCAGAGGACGACAGCATCG 315
QY 548 TGAGGTGAATGGCCGAGTACCTGGATGGAGAGAGACTTTAAACCCAAATTGCAGGGTCCT 607
Db 316 TGAGGTGAACGGCCGCTGTATCGGAAGGGGAGACCCCTTCAGCCCCACTGCAGCATCCG 375
QY 608 GTGCCGCTGTGATGACGGTGGCTTCACC 635
Db 376 CTGCCGCTGCGAGGACGGCGGCTTCACC 403

RESULT 15
BQ073722
LOCUS
DEFINITION BQ073722 1073 bp mRNA linear EST 02-APR-2002
AGENCOURT_7046577 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:5806602
5', mRNA sequence.
ACCESSION BQ073722
VERSION BQ073722.1 GI:19902768
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1073)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2047 row: k column: 19
High quality sequence stop: 517.
FEATURES
Location/Qualifiers
source
1. .1073
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5806602"
/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 166 a 384 c 347 g 176 t
ORIGIN
Query Match 40.3%; Score 256; DB 14; Length 1073;
Best Local Similarity 81.9%; Pred. No. 1.4e-55;
Matches 295; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 276 CTGCCACTTCCTTCCTCGCCTTCTCTCAATGGTGTGTGCCAGCTGTGCCGACACC 335
```

Db 2 CTGGCCTTCTCCCTCCTCTGCTCTCAAGGTGCGTACCCAGTGTGCCGACACCA 61
QY 336 TGTACCTGTCTTGGACACCAACCCAGTGCACAGGGGTACCCCTGGTGTGGATGGC 395
Db 62 TGTACCTGCCCTGGCCACCTCCCGGATGCCCGCTGGGAGTACCCCTGGTGTGGATGGC 121
QY 396 TGTGGCTGTGTAAAGTGTGTGCACGAGGCTGGGGGAGTCCCTGGACCAACCTGCATGTC 455
Db 122 TGTGGCTGTGCCGGGTATGTGCACGGCGGCTGGGGGAGCCCTGGGACCAACTCCACGTC 181
QY 456 TGGACCCAGCCAGGGCCCTGGTTTGTACGCTGGGGCAGGCCCTGGCGGCCATGGGGCT 515
Db 182 TGGACGCCAGCCAGGGCCCTGGTCTGCCAGCCCCGGGGCAGGACCCGGTGGCCGGGGGCC 241
QY 516 GTGTGTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGAGGTACCTGGAT 575
Db 242 CTGTGCTCTTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGGCCGCTGTATCGGGAA 301
QY 576 GGAGAGACCTTTAAACCCCAATTGCAGGTCCTGTGCCGCTGTGATGACGGTGGCTTCACC 635
Db 302 GGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGACGGGGCTTCACC 361

Search completed: July 29, 2003, 02:51:28
Job time : 1128.2 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 28, 2003, 15:59:01 ; Search time 2454.08 Seconds
(without alignments)
10471.464 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_883
Perfect score: 883
Sequence: 1 GACGCTCTGATCTCCAGAG.....ACCACCTGTGGGCTGGGCAT 883

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	821.4	93.0	1741	10	AF259981	AF259981 Rattus no
2	727.6	82.4	1734	6	AR210324	AR210324 Sequence
3	727.6	82.4	1734	6	AR210325	AR210325 Sequence
4	727.6	82.4	1734	10	AF100778	AF100778 Mus muscu
5	718.2	81.3	1739	10	AF126063	AF126063 Mus muscu
6	434.6	49.2	1266	6	AX076919	AX076919 Sequence
7	434.6	49.2	1266	6	AX464186	AX464186 Sequence
8	434.6	49.2	1283	9	AF083500	AF083500 Homo sapi
9	434.6	49.2	1293	6	AR210322	AR210322 Sequence
10	434.6	49.2	1293	6	AR210323	AR210323 Sequence
11	434.6	49.2	1309	9	AF074604	AF074604 Homo sapi
12	434.6	49.2	1427	9	AF100780	AF100780 Homo sapi
13	434.6	49.2	1450	9	BC017782	BC017782 Homo sapi
14	423	47.9	738	6	AR210337	AR210337 Sequence
15	420.6	47.6	841	6	AR210338	AR210338 Sequence
16	256.6	29.1	137964	2	AC126895	AC126895 Rattus no
17	221.6	25.1	137964	2	AC126895	AC126895 Rattus no
18	221.6	25.1	226303	2	AC095418	AC095418 Rattus no
19	216.2	24.5	61072	10	AL731698	AL731698 Mouse DNA
20	199.2	22.6	216757	2	AL669906	AL669906 Mus muscu
21	169.8	19.2	107260	9	AL139352	AL139352 Human DNA
22	138.8	15.7	1805	5	CHKCEF	J04496 Chicken CEF
23	128.6	14.6	1585	5	NVI271167	AJ271167 Notophtha
24	128.4	14.5	2075	6	AR018957	AR018957 Sequence
25	128.4	14.5	2075	6	AR118595	AR118595 Sequence
26	128.4	14.5	2075	6	AR130891	AR130891 Sequence
27	128.4	14.5	2075	6	AX206708	AX206708 Sequence
28	128.4	14.5	2075	6	I11636	I11636 Sequence 1
29	128.4	14.5	2075	6	I32210	I32210 Sequence 1
30	128.4	14.5	2075	9	HUMCONGRO	M92934 Homo sapien
31	128.4	14.5	2312	9	HSCCTGF	X78947 H.sapiens m
32	128.4	14.5	2998	6	AR119211	AR119211 Sequence
33	128.4	14.5	2998	6	ARI51276	ARI51276 Sequence
34	127.4	14.4	1975	6	A28405	A28405 Sequence I
35	127.4	14.4	1975	6	A28444	A28444 nov mRNA se
36	127.4	14.4	1976	5	GGNOVMRNA	X59284 G.gallus no
37	126.6	14.3	2052	9	HSCYR61	Y11307 H.sapiens C
38	125.8	14.2	2338	6	E37595	E37595 Monoclonal
39	125.8	14.2	2338	10	AB023068	AB023068 Rattus no
40	125.8	14.2	2345	6	AX402019	AX402019 Sequence
41	125.8	14.2	2345	10	AF120275	AF120275 Rattus no
42	125.8	14.2	2350	6	ARI94011	ARI94011 Sequence
43	125	14.2	1146	6	E13814	E13814 cDNA encodi
44	125	14.2	1887	9	AF031385	AF031385 Homo sapi
45	125	14.2	1935	9	AF003594	AF003594 Homo sapi

ALIGNMENTS

RESULT 1
AF259981
LOCUS AF259981 1741 bp mRNA linear ROD 09-MAY-2000
DEFINITION Rattus norvegicus CCN family protein COP-1 (Cop-1) mRNA, complete cds.
ACCESSION AF259981
VERSION AF259981.1 GI:7739780
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1741)
AUTHORS Zhang,R., Averboukh,L., Zhu,W., Zhang,H., Jo,H., Dempsey,P.J.,

Coffey,R.J., Pardee,A.B. and Liang,P.
Identification of rCop-1, a new member of the CCN protein family,
as a negative regulator for cell transformation
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)
84114629
9742130
REFERENCE 2 (bases 1 to 1741)
AUTHORS Liang,P.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer
Center, 649 MRB II, Nashville, TN 37232, USA
FEATURES
source Location/Qualifiers
1..1741
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1..1741
/gene="Cop-1"
262..1014
/gene="Cop-1"
/feature="secreted protein"
/codon_start=1
/product="CCN family protein COP-1"
/protein_id="AAF69011.1"
/db_xref="GI:7739781"
/translation="MRGSLIRLLATSFLCLLSMVCAQLCRTPCTPWTPPQCPQGV
LVLDGCGCKVCARRLTESCEHLHVCEPSQGLVCPQGAGPGGHGAVCLLDEDDGDCEV
NGRRYLDGETFKPNCRVLCRCDGGFTCLPLCSEDTVLPWDPCRPKRIOVPGKCCPE
WVCDQVTPAIQRSAAQHQLSALVTPASADAPWPNWSTAWGPCSTTCGLGIATRVSN
QNRFOLEIQRLCLPRPCLAARSHSSWNSAF"
BASE COUNT 386 a 491 c 480 g 384 t
ORIGIN
Query Match 93.0%; Score 821.4; DB 10; Length 1741;
Best Local Similarity 97.8%; Pred. No. 2.1e-188;
Matches 864; Conservative 0; Mismatches 16; Indels 3; Gaps 3;
QY 1 GACGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGCCCTTGCAAGGCTGCAGCC 60
|||||
Db 17 GACGCTTCTGATCTCCAGAGGACCCTGGGTGGGACAGGGG-CCTTGGCAAGGCTGCAG-C 74
QY 61 GCTGGGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 120
|||||
Db 75 GCTGGGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 134
QY 121 CTGTACGCTTGCTCTAAAGCTTTAGCACTTGTTGGTGGCTTGGGCTTCACACACTGTCTAGA 180
|||||
Db 135 CTGTACGCTTGCTCTAAACCCTTAGCACTTGTTGGTGGCTTGGGCTTCACACACTGTCTAGA 194
QY 181 CACCTTCGTGTGGCCCTCCACGGCCTCACTTCAAGCTTTGAAGCTGGCTCCACAAGGGAC 240
|||||
Db 195 CACCTTCGTGTGGCCCTCCACGG-CTCACCTCCAGGTTTGAAGCTGGCTCCACAAGGGAC 253
QY 241 ACGGTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCCTTCTGCTTC 300
|||||
Db 254 ACGGTGACATGAGGGGCAGCCCACTGATCCGTCTTCTGGCCACTTCCCTTCTGCTTC 313
QY 301 TCTCAATGGTGTGTGCCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCAACCC 360
Db 314 TCTCAATGGTGTGTGCCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCAACCC 373
QY 361 AGTCCCCACAGGGGTACCCCTGGTGGATGGCTGTGGCTGTAAAGTGTGTGCAC 420
Db 374 AGTCCCCACAGGGGTACCCCTGGTGGATGGCTGTGGCTGTAAAGTGTGTGCAC 433
QY 421 GGAGGCTGGGGAGTCTCGGACCACTGCATGTCTCGACCCAGCCAGGGCCTGGTTT 480
Db 434 GGAGGCTGACGGAGTCTCGGAACACCTGCATGTCTCGAACCCAGCCAGGGCCTGGTTT 493
QY 481 GTCAGCCTGGGCGAGCCCTGGCGGCATGGGGCTGTGTGTCTCTTGGATGAGGATGACG 540
Db 494 GTCAGCCTGGGCGAGCCCTGGCGGCATGGGGCTGTGTGTCTCTTGGATGAGGATGACG 553
QY 541 GTAGCTGTAGGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCA 600

|||||
Db 554 GTGACTGTGAGTGAATGGCCGAGGTACTGGATGGAGAGACCTTTAAACCCCAATTGCA 613
QY 601 GGGTCCCTGTGCGCTGTGATGACGGTGGCTTTCACCTGCCTGCCGTGTCAGTGAGGATG 660
|||||
Db 614 GGGTCCCTGTGCGCTGTGATGACGGTGGCTTTCACCTGCCGTGCCGTGTCAGTGAGGATG 673
QY 661 TGGCGCTGCCAGCTGGGACTGCCCCACGCCCCAAGAGAATACAGGTGCCAGGAAAGTGCT 720
Db 674 TGACGCTGCCAGCTGGGACTGCCCCACGCCCCAAGAGAATACAGGTGCCAGGAAAGTGCT 733
QY 721 GCCCCGAGTGGGTATGTGACCCAGGGAGTGACACCCGGCGATCCAGCGCTCCACGGCGCAAG 780
Db 734 GCCCCGAGTGGGTATGTGACCCAGGGAGTGACACCCGGCGATCCAGCGCTCCAGCGCGCAAG 793
QY 781 GACACCAACTTCTGCCCTTGTGCTCACTCCTGCCTCTGCTGATGCTCCTTGTCCAAATTGGA 840
Db 794 GACACCAACTTCTGCCCTTGTGCTCACTCCTGCCTCTGCTGATGCTCCTTGTCCAAATTGGA 853
QY 841 GCACAGCCTGGGCCCCCTGCTCAACCACTGCTGGGCTGGGCAT 883
Db 854 GCACAGCCTGGGCCCCCTGCTCAACCACTGCTGGGCTGGGCAT 896
RESULT 2
AR210324
LOCUS AR210324 1734 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 17 from patent US 6387657.
ACCESSION AR210324
VERSION AR210324.1 GI:21512525
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 17 14-MAY-2002;
FEATURES Location/Qualifiers
source 1..1734
BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN
Query Match 82.4%; Score 727.6; DB 6; Length 1734;
Best Local Similarity 91.9%; Pred. No. 1.1e-165;
Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;
QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGTGGGACAGGGGCCCTTGCAAGGCTGCAGCCGC 62
|||||
Db 13 CGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCCCTTGCGAGGCTGCAGCTGC 72
QY 63 TG-GGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAACCTGAGAGCTAAGAGGCTCC 121
|||||
Db 73 TGTGGCAGTAGCTTGGGATGGAGGCTTTCTTGTCTGGGAACCTGAGAGGCTGAGAGGCTCC 132
QY 122 TGTGAG---CTTGTCTCTAAAGCTTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCA 178
|||||
Db 133 TGTGAGGCTCCTGTCTCTAAACTCTTGGCACTTGGGTGGCTTGGGCTTCACACACTGTCA 192
QY 179 GACACCTTGTGGTGGCTTCCACGGCCTCACCTTTCAGGTTTGAAGCTTGGCTCCACAAGGG 238
Db 193 GACACCTTCTTGGTGGCTTCTCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 246
QY 239 ACACGGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTC 298
Db 247 ACACGGTGACATGAGGGGCGAACCCCACTGATCCATCTTCTGGCCATTTCTCTCTGCTC 306
QY 299 TCTCTCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACTGTCTTGGACACCACC 358
Db 307 TCTCTCAATGGTGTATTCCCAGCTGTGCCCGCAGCACCCCTGTGCCTGTCTTGGACACCACC 366

QY	359	CCAGTCCCCACAGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGCTGTAAAGTGTGTC	418
Db	367	CCAGTCCCCACCGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGCTGTGAGTGTGTC	426
QY	419	ACGGAGGCTGGGGAGTCCCTGCGACCACTGCATGTCTGCGACCCAGCCAGGCGCTGGT	478
Db	427	ACGGAGGCTGGGGAGTCCCTGCGACCACTGCATGTCTGCGACCCAGCCAGGCGCTGGT	486
QY	479	TTGTACGCTGGGGCAGGCCCTGGCGGCATGGGGCTGTGTGCTCTTGGATGAGGATGA	538
Db	487	TTGTACGCTGGGGCAGGCCCTAGTGGCGGTGTGTGCTCTTGAAGAGGATGA	546
QY	539	CGGTAGCTGTGAGGTGAATGGCCGACGTACCTGGATGGAGAGACCTTTAAACCCCAATTG	598
Db	547	CGGGAGCTGTGAGGTGAATGGCCGACGTACCTGGATGGGAGACCTTTAAACCCCAATTG	606
QY	599	CAGGGTCCTGTGCCGCTGTGATGACGGTGGCTTCACCTGCCCTGCCCTGTGCAGTGAGGA	658
Db	607	CAGGGTTTGTGCCGCTGTGATGACGGTGGTTTCACCTGCCCTGCCCTGTGCAGTGAGGA	666
QY	659	TGTGCGGCTGCCCAGCTGGGACTGCCCCAGCCCCAAGAGAATACAGGTGCCAGGAAAGTG	718
Db	667	TGTGCGGCTGCCCAGCTGGGACTGCCCCAGCCCCAGGAGAATACAGGTGCCAGGAAAGTG	726
QY	719	CTGCCCCGAGTGGGTATGTGACCAGGGAGTGA---CACCGGGATCCAGCGCTCCACGGC	775
Db	727	CTGCCCCGAGTGGGTGTGTGACCAGGCAGTGTGACGCGGCAATCCAGCCCTCCTCAGC	786
QY	776	GCAAGGACACCAACTTCTGCCCTTGTCACTCCTCCTCTGTGATGCTCTTGTCCAAA	835
Db	787	CCAAGGACACCAACTTCTGCCCTTGTCACTCCTCCTGTGATGCTGCGGCTGTCCAAA	846
QY	836	TTGGAGCACAGCCTGGGGCCCCCTGCTCAACACACCTGTGGCTGGGCAT	883
Db	847	CTGGAGCACAGCCTGGGGCCCCCTGCTCAACACACCTGTGGTTGGGCAT	894
RESULT 3			
AR210325/c			
LOCUS	AR210325	1734 bp	DNA linear PAT 20-JUN-2002
DEFINITION	Sequence 18 from patent US 6387657.		
ACCESSION	AR210325		
VERSION	AR210325.1 GI:21512526		
KEYWORDS	WISP polypeptides and nucleic acids encoding same		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1734)		
AUTHORS	Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.		
TITLE	WISP polypeptides and nucleic acids encoding same		
JOURNAL	Patent: US 6387657-A 18 14-MAY-2002;		
FEATURES	Location/Qualifiers		
source	1..1734		
BASE COUNT	393 a	495 c	491 g 355 t
ORIGIN	/organism="unknown"		
Query Match	82.4%	Score 727.6;	DB 6; Length 1734;
Best Local Similarity	91.9%;	Pred. No. 1.1e-165;	
Matches 816;	Conservative 0;	Mismatches 59;	Indels 13; Gaps 4;
QY	3	CGCTTCTGATCTCCAGAGGACCCCTGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGC	62
Db	1722	CGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCCCTTGGCGAGGCTGCAGCTGC	1663
QY	63	TG-GGCAGTGGCTTGAATGGAGGTCTTTATTAAGTGGAACTGAGGAGCTAAGAGGCTCC	121
Db	1662	TGTGGCAGTAGCTTGGGATGGAGGTCTTTCTGTCTGGAACTGAGGAGCTGAGAGGCTCC	1603
QY	122	TGTCAG---CTTGCTCTAAAGTCTTAGCACTGTGGTGGCTTGGGCTTCACACACTGTCA	178
Db	1602	TGTCAGGCTCCTGTCTCTAAACTCTTGGCACTTGGGTGGCTTGGGCTTCACACACTGTCA	1543

QY	179	GACACCTTCGTGGTGGCTCCACGGCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG	238
Db	1542	GACACCTTCTTGGTGGCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG	1489
QY	239	ACACGGTGACATGAGGGGACGCCCACTGATCCATCTTCTGGCCACTTCCTTCTCTGCGCT	298
Db	1488	ACACGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCATTTCTTCTCTGCGAT	1429
QY	299	TCTCTCAATGGTGTGTGCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACC	358
Db	1428	TCTCTCAATGGTGTATTCCACAGCTGTGCCACGACCCCTGTGCCTGTCTTGGACACCAACC	1369
QY	359	CCAGTGCCACACAGGGGTACCCCTGGTGGTGGATGGATGGCTGTGCTCTTGGATGAGGATGA	418
Db	1368	CCAGTGCCACACCGGGGTACCCCTGGTGGTGGATGGATGGCTGTGCTCTTGGATGAGGATGA	1309
QY	419	ACGGAGGCTGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGCGCTGGT	478
Db	1308	ACGGAGGCTGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGCGCTGGT	1249
QY	479	TTGTACGCTGGGGCAGGCCCTGGCGCCCATGGGCTGTGTGCTCTTGGATGAGGATGA	538
Db	1248	TTGTACGCTGGGGCAGGCCCTGAGTGGCGGTGTGTGCTCTTGGATGAGGATGA	1189
QY	539	CGGTAGCTGTGAGGTGAATGGCCGACGTACCTGGATGGAGAGACCTTTAAACCCCAATTG	598
Db	1188	CGGGAGCTGTGAGGTGAATGGCCGACGTACCTGGATGGGAGACCTTTAAACCCCAATTG	1129
QY	599	CAGGGTCCTGTGCCGCTGTGATGACGGTGGCTTCACCTGCCCTGCCGCTGTGCAGTGAGGA	658
Db	1128	CAGGGTTTGTGCCGCTGTGATGACGGTGGTTTACCTGCCCTGCCGCTGTGCAGTGAGGA	1069
QY	659	TGTGCGGCTGCCCAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGCCAGGAAAGTG	718
Db	1068	TGTGCGGCTGCCCAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGCCAGGAAAGTG	1009
QY	719	CTGCCCCGAGTGGGTATGTGACCAGGGAGTGA---CACCGGGATCCAGCGCTCCACGGC	775
Db	1008	CTGCCCCGAGTGGGTGTGTGACCAGGCAGTGTGATGACCGGCAATCCAGCCCTCCTCAGC	949
QY	776	GCAAGGACACCAACTTTCTGCCCTTGTCACTCCTCCTCCTGTGATGCTCTTGTCCAAA	835
Db	948	CCAAGGACACCAACTTTCTGCCCTTGTCACTCCTCCTCCTGTGATGCTGCCCTGTCCAAA	889
QY	836	TTGGAGCACAGCCTGGGGCCCCCTGCTCAACACACCTGTGGGCTGGGCAT	883
Db	888	CTGGAGCACAGCCTGGGGCCCCCTGCTCAACACACCTGTGGGTTGGGCAT	841
RESULT 4			
AF100778			
LOCUS	AF100778	1734 bp	mRNA linear ROD 17-DEC-1998
DEFINITION	Mus musculus connective tissue growth factor related protein WISP-2 (Wisp2) mRNA, complete cds.		
ACCESSION	AF100778		
VERSION	AF100778.1 GI:4028578		
KEYWORDS	Mus musculus.		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1734)		
REFERENCE	Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A., Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C., Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D., Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.		
AUTHORS	WISP genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors		
TITLE	Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)		
JOURNAL	99061933		
MEDLINE	9843955		
PUBMED			

Db126TGCCCCGTGGGAGTACCCCTGGTGGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG185

QY423AGGCTGGGGAGTCCTGCGACCAACCTGCATGTCTGCGACCCCCAGCCAGGGCCTGGTTGT482

Db186CGGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGC245

QY483CAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTTGGATGAGGATGACGGT542

Db246CAGCCCGGGCAGGACCCGGTGGCCGGGGCCCTGTGCCTCTTTGGCAGAGGACGACAGC305

QY543AGCTGTAGGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG602

Db306AGCTGTAGGTGAACGGCCCGCCTGTATCGGGAGAGGGGAGACCTTCCAGCCCCACTGCAGC365

QY603GTCCTGTGCCGCTGTGATGACGGTGGCTTCACTGCCCTGCCGCTGTGCAGTGAGGATGTG662

Db366ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACTGCGTGCCGCTGTGCAGCGAGGATGTG425

QY663CGGCTGCCCAGCTGGGACTGCCCCACGCCCCCAAGAGATAACAGGTGCCAGGAAAGTGTGC722

Db426CGGCTGCCCAGCTGGGACTGCCCCACCCCCAGGAGGTCGAGGTCTTGGGCAAGTGTGC485

QY723CCCGAGTGGGTATGTGACCAAGGAGTGACACCGGCGATCCAGCGCTCCAGGGCGCAAGGA782

Db486CCTGAGTGGGTGTGCGGCCAAGAGGGGGACTGGGGACCAAGCCCTTCCA-GCCCCAAGGA543

QY783CACCAACTTTCTGCCCTTGTCACTCCTGCTGCTGATGCTCCTTGTCCAAATTTGGAGC842

Db544CCCCAGTTTTCTGGCCTTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCCAGAATGGAGC603

QY843ACAGCCTGGGGCCCTGCTCAACCAACCTGTGGGCTGGGCAT883

Db604ACGGCCTGGGGACCCCTGCTCGACCACCTGTGGGCTGGGCAT644

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2003, 09:03:08 ; Search time 233.823 Seconds
(without alignments)
8504.365 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_883

Perfect score: 883

Sequence: 1 GACGGTCTGTGATCTCCAGAG.....ACCACCTGTGGGCTGGGCAT 883

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	883	100.0	1708	AAZ07516	Rat HICP polypepti
2	727.6	82.4	1734	AAAX76488	Mouse WISP-2 prote
3	635	71.9	753	AAZ07517	Rat HICP polypepti
4	566	64.1	681	AAZ07521	Rat HICP mature po
5	553.2	62.7	753	AAAX76489	Mouse WISP-2 prote
6	460.6	52.2	1522	AAAX16595	Human growth facto
7	440	49.8	1337	AAH46952	Human secreted pro
8	439.8	49.8	1352	AAH46936	Human secreted pro
9	434.6	49.2	1266	AAA30048	Human PRO261 nucle

10	434.6	49.2	1266	22	AAZ07516	Human cDNA sequenc
11	434.6	49.2	1266	22	AAAF60368	PRO261 coding sequ
12	434.6	49.2	1266	22	AAC97451	Human angiogenesis
13	434.6	49.2	1267	21	AAAX77566	Human PRO261 cDNA
14	434.6	49.2	1285	19	AAV29260	Human connective t
15	434.6	49.2	1293	20	AAAX76486	Human WISP-2 prote
16	434.6	49.2	1309	22	AAH28214	Connective tissue
17	430.2	48.7	750	20	AAAX76487	Human WISP-2 prote
18	430.2	48.7	1257	20	AAAX28435	EGF-like homologue
19	423	47.9	738	20	AAAX76501	Human WISP-2 prote
20	420.6	47.6	841	20	AAAX76502	Human WISP-2 prote
21	210	23.8	210	20	AAZ07519	Rat HICP IGFBP dom
22	203.4	23.0	2136	22	AAK94706	Human full-length
23	177	20.0	177	20	AAZ07518	Rat HICP VWC domai
24	169.8	19.2	586	22	ABA59703	Human foetal liver
25	169.8	19.2	586	22	ABA28232	Probe #6698 for ge
26	169.8	19.2	586	22	AAK07972	Human brain expres
27	169.8	19.2	586	22	AAK33842	Human bone marrow
28	169.8	19.2	586	22	AAI16552	Probe #6485 for ge
29	169.8	19.2	586	22	AAI39568	Probe #8254 used t
30	169.8	19.2	586	24	ABS08689	Human genome-deriv
31	169.8	19.2	13255	22	AAK76842	Human immune/haema
32	139.8	15.8	199	22	ABA72245	Human foetal liver
33	139.8	15.8	199	22	ABA38112	Probe #16578 for g
34	139.8	15.8	199	22	AAK20667	Human brain expres
35	139.8	15.8	199	22	AAK46811	Human bone marrow
36	139.8	15.8	199	22	AAI25749	Probe #15682 for g
37	139.8	15.8	199	22	AAI52650	Probe #21336 used
38	139.8	15.8	199	24	ABS21124	Human genome-deriv
39	128.4	14.5	2075	16	AAAT04226	Connective tissue
40	128.4	14.5	2075	18	AAAT45360	Human connective t
41	128.4	14.5	2075	18	AAAT51234	Connective tissue
42	128.4	14.5	2075	19	AAV38085	Human connective t
43	128.4	14.5	2075	20	AAAX61317	Human connective t
44	128.4	14.5	2075	21	AAA11278	Human connective t
45	128.4	14.5	2075	21	AAA11280	Human connective t

ALIGNMENTS

RESULT 1

AAZ07516
ID AAZ07516 standard; cDNA; 1708 BP.

XX
AC AAZ07516;

XX
DT 26-NOV-1999 (first entry)

XX
DE Rat HICP polypeptide encoding cDNA.

XX
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.

XX
OS Rattus sp.

XX
PN WO9947556-A2.

XX
PD 23-SEP-1999.

XX
PF 18-MAR-1999; 99WO-US05999.

XX
PR 19-MAR-1998; 98US-0044273.

XX
(TUFT) TUFTS COLLEGE.

XX
Castellot JJ;

XX
WPI; 1999-562060/47.

XX
P-PSDB; AAY27434.

XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications -

Db 421 CCCAGCTGGGACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTGCTGCCCCCGAG 480
QY 729 TGGGTATGTGACCAAGGAGTGACACCGCGCGATCCAGCGCTCCACGGCGCAAGGACACCAA 788
Db 481 TGGGTATGTGACCAAGGAGTGACACCGCGCGATCCAGCGCTCCACGGCGCAAGGACACCAA 540
QY 789 CTTTCTGCCCTTGTCACTCCTGCTCTGCTGATGCTCCTTGTCCTCAAAATTGGAGCACAGCC 848
Db 541 CTTTCTGCCCTTGTCACTCCTGCTCTGCTGATGCTCCTTGTCCTCAAAATTGGAGCACAGCC 600
QY 849 TGGGGCCCTGCTCAACCACCTGTGGGCTGGGCAT 883
Db 601 TGGGGCCCTGCTCAACCACCTGTGGGCTGGGCAT 635

RESULT 4
AAZ07521
ID AAZ07521 standard; cDNA; 681 BP.
XX
AC AAZ07521;
XX
DT 26-NOV-1999 (first entry)
XX
DE Rat HICP mature polypeptide coding sequence.
XX
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX
OS Rattus sp.
XX
PN W09947556-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Castellet JJ;
XX
DR WPI; 1999-562060/47.
DR P-PSDB; AAY27440.
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
used in methods to identify modulators or in diagnostic applications -
XX
PS Disclosure; Fig 2; 108pp; English.

XX
CC The invention provides a rat heparin-induced CCN-like protein (HICP)
protein. Agents that stimulate or inhibit HICP protein activity or
expression, antisense HICP nucleic acid molecules and HICP antibodies,
can be used to modulate cell-associated activity. HICP modulators can be
used to treat disorders characterized by aberrant HICP protein activity
or expression. Probes capable of hybridizing to HICP mRNA or antibodies
specific for HICP can be used to detect HICP activity in a biological
sample. HICP can be used to treat disorders, such as a cardiovascular or
fibrotic disorder, characterized by aberrant cell proliferation. The
present sequence represents the coding sequence of rat HICP mature
polypeptide.
XX
SQ Sequence 681 BP; 121 A; 211 C; 215 G; 134 T; 0 other;

Query Match 64.1%; Score 566; DB 20; Length 681;
Best Local Similarity 100.0%; Pred. No. 4.5e-147;
Matches 566; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCAGTCCCCACAGGGGTA 377
Db 1 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCAGTCCCCACAGGGGTA 60
QY 378 CCCCTGGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGCACGGAGGCTGGGGGAGTCC 437

Db 61 CCCCTGGTGTGGATGGCTGTGGCTGTAAAGTGTGTGCAGGAGGCTGGGGAGTCC 120
QY 438 TCGGACACCTGCATGTCTGCGACCCAGCCAGGGCCTGGTTGTTCAGCCTGGGGCAGGC 497
Db 121 TCGGACACCTGCATGTCTGCGACCCAGCCAGGGCCTGGTTGTTCAGCCTGGGGCAGGC 180
QY 498 CCTGCGGCCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT 557
Db 181 CCTGCGGCCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT 240
QY 558 GGCCGCGAGGTACCTGGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTCTGCGCGTGT 617
Db 241 GGCCGCGAGGTACCTGGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTCTGCGCGTGT 300
QY 618 GATGACGGTGGCTTCACCTGCCCTGCCGCTGTGCAGTGAGGATGTGCGGCTGCCACAGTGG 677
Db 301 GATGACGGTGGCTTCACCTGCCCTGCCGCTGTGCAGTGAGGATGTGCGGCTGCCACAGTGG 360
QY 678 GACTGCCACACGCCCAAGAGAAATACAGGTGCCAGGAAAGTGCTGCCCGAGTGGGTATGT 737
Db 361 GACTGCCACACGCCCAAGAGAAATACAGGTGCCAGGAAAGTGCTGCCCGAGTGGGTATGT 420
QY 738 GACCAGGGAGTGAACACCGCGGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCC 797
Db 421 GACCAGGGAGTGAACACCGCGGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCC 480
QY 798 CTTGTCACTCCTGCTCTGCTGATGCTCCTTGTCCTCAAAATTGGAGCACAGCCCTGGGGCCCC 857
Db 481 CTTGTCACTCCTGCTCTGCTGATGCTCCTTGTCCTCAAAATTGGAGCACAGCCCTGGGGCCCC 540
QY 858 TGCTCAACACACCTGTGGGCTGGGCAT 883
Db 541 TGCTCAACACACCTGTGGGCTGGGCAT 566

RESULT 5
AAX76489/c
ID AAX76489 standard; DNA; 753 BP.
XX
AC AAX76489;
XX
DT 06-AUG-1999 (first entry)
XX
DE Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
connective tissue growth factor; cancer; melanoma; arteriosclerosis;
leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
connective tissue disorder; catabolic state; inflammation;
testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Mus sp.
XX
PN W09921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI; 1999-337420/28.

Db 242 GCTCTGCAGG-----GACATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTC 293

QY 287 CTTCCTCTGCTTCTCTCAATGGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCC 346

Db 294 CCTCCTCTGCTCCTCTCAAAGGTGGTACCCAGCTGTGCCCGACACCATGTACCTGCC 353

QY 347 TTGGACACACCCAGTGCCACAGGGGGTACCCCTGGTGTGATGGCTGTGGCTGCTG 406

Db 354 CTGGCCACTCCCGATGCCCGCTGGGAGTACCCCTGGTGTGATGGCTGTGGCTGCTG 413

QY 407 TAAAGTGTGTGCACGGAGGCTGGGGAGTCTCGGACCACTTCGATGTCTCGACCCCG 466

Db 414 CCGGGTATGTGCACGGCGGCTGGGGAGCCCTCGGACCAACTCCAGTCTCGACGCCAG 473

QY 467 CCAGGGCCTGGTTGTTCAGCCTGGGCGAGGCCCTGTGGCGCCATGGGGCTGTGTCTCT 526

Db 474 CCAGGGCCTGGTCTGCCAGCCCGGGCAGGACCCGGTGGCCGGGGCCCTGTGCTCTCT 533

QY 527 GGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCCGCAGGTACCTGGATGGAGACCT 586

Db 534 GGCAGAGGACGACAGCTGTGAGGTGAACGGCCCTGTATCGGAAGGGAGACCT 593

QY 587 TAAACCCCAATTGCAGGGTCTCTGTCGGCTGTGATGACGGTGGCTTCACCTGCCCTGCC 646

Db 594 CCAGCCCACTGCAGCATCCGCTGCCGCTGCAGGACGGCGGCTTCACCTGCCCTGCCCT 653

QY 647 GTGAGTGAGGATGTGCGGCTGCCAGCTGGGACTGCCACGCGCCCAAGAGAATACAGGT 706

Db 654 GTGACGAGGATGTGCGGCTGCCAGCTGGGACTGCCCCACCCAGGAGGTGCGAGGT 713

QY 707 GCCAGGAAGTGTGCCCCGAGTGGGTATGTACCGAGGGAGTGACACCGCGATCCAGCG 766

Db 714 CTTGGCAAGTGTGCCCTGAGTGGGTGTGCGGCCAAGGAGGGGACTGGGACCCAGCC 773

QY 767 CTCCACGGCGCAAGGACACCAACTTCTGCCCTTGTCACTCTCCTCTCTGATGCTCC 826

Db 774 CTTCCAGCCCAAGGACCCAGTTTCTGGCCTTGTCTCTCCCTGCCCTGGTGTCCC 833

QY 827 TTGTCCAAATTGGAGCACAGCCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCAT 883

Db 834 CTGCCCAAGATGGAGCACGGCCTGGGACCCCTGCTCGACCACTGTGGGCTGGGCAT 890

RESULT 7

AAH46952

ID AAH46952 standard; cDNA; 1337 BP.

XX

AC AAH46952;

XX

DT 25-SEP-2001 (first entry)

XX

DE Human secreted protein encoding cDNA (clone Id HB0DE48).

XX

KW Secreted protein; immunosuppressive; antiarthritic; antirheumatic;

KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;

KW nootropic; neuroprotective; antibacterial; virucide; fungicide; human;

XX opthalmological; gene therapy; ss.

OS Homo sapiens.

XX

PN WO200155430-A1.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US01431.

XX

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 12-SEP-2000; 2000US-0231968.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;

PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;

PI Ni J, Ruben SM, Barash SC;

XX

DR WPI; 2001-476220/51.

DR P-PSDB; AAB85542.

XX

PT 17 isolated nucleic acid molecules encoding human secreted proteins,

PT used to preventing, treating or ameliorating a medical condition -

XX

PS Claim 1; Page 427-428; 482pp; English.

XX

CC The invention provides novel human secreted proteins and polynucleotides

CC encoding them. The secreted proteins can be expressed by standard

CC recombinant methodology. The secreted proteins and polynucleotides are

CC used to prevent, treat or ameliorate a medical condition in e.g. humans,

CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can

CC also be used in diagnosing a pathological condition. The antibodies to

CC the proteins can also be used in alleviating symptoms associated with the

CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme

CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or

CC treated include autoimmune diseases e.g. rheumatoid arthritis,

CC hyperproliferative disorders e.g. neoplasms of the breast or liver,

CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders

CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.

CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and

CC ocular disorders e.g. corneal infection. The polypeptides can also be

CC used to aid wound healing and epithelial cell proliferation, to prevent

CC skin aging due to sunburn, to maintain organs before transplantation, for

CC supporting cell culture of primary tissues, to regenerate tissues and in

CC chemotaxis. The polypeptides can also be used as a food additive or

CC preservative to increase or decrease storage capabilities. The present

CC sequence represents a human secreted protein encoding cDNA.

XX

SQ Sequence 1337 BP; 257 A; 427 C; 396 G; 252 T; 5 other;

Query Match 49.8%; Score 440; DB 22; Length 1337;

Best Local Similarity 78.3%; Pred. No. 5.2e-112;

Matches 539; Conservative 3; Mismatches 138; Indels 8; Gaps 1;

QY 196 CTCACGGCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAAGGACACGGTGACATGAGG 255

Db 2 CTTACAGTTTCACTTCAGGCTCAAARCTGGSTCTGCA-----GGGACATGAGAG 53

QY 256 GCAGCCCACTGATCCATCTTCTGGCCACTTCTCCTCTCCCTCTCTCAATGGTGTG 315

Db 54 GCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTCTCAAGGTGCGTA 113

QY 316 CCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCCACCCAGTGCCACAGGGG 375

Db 114 CCCAGCTGTGCGGACACCCATGTACCTGCCCTGCCCTGCCGATGCCCGTGGGAG 173

QY 376 TACCCCTGGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACGGAGGCTGGGGAGT 435

Db 174 TACCCCTGGTGTGGATGGCTGTGGCTGTGCCGGTATGTGACGGCGGCTGGGGAGC 233

QY 436 CTTGCGACCACTGCATGTCTGCGACCCCGCAGCCAGGCTGTGTGAGCCTGGGGCAG 495

Db 234 CTTGCGACCACTGCACGTCTGCGACCGCCAGCCAGGCTGTGTGCCAGCCCGGGCAG 293

QY 496 GCCCTGGCGGCCATGGGGCTGTGTGTCTTGGATGAGGATGACGGTAGCTGTGAGGTGA 555

Db 294 GACCCGGTGMCGGGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCTGTGAGGTGA 353

QY 556 ATGGCCCGCAGGTAGCTGGATGGAGAGACCTTTAAACCCCAATTCAGGGTCTGTGCCGCT 615

Db 354 ACGCCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCCACTGCAGCATCCGCTGCCGCT 413

QY 616 GTGATGACGGTGGCTTCACTGCCTGCCGCTGTGCGAGTGAGGATGTGCGGCTGCCAGCT 675

Db 414 GCGAGGACGGGGTTCACTGCTGCCGCTGTGCGAGGAGGATGTGCGGCTGCCAGCT 473

QY 676 GGGACTGCCACGGCCCAAGAAATACAGGTGCCAGGAAGTGTGCCCGAGTGGGTAT 735

KW antibody dependent enzyme mediated prodrug therapy; ss.
OS Homo sapiens.
XX WO200015666-A2.
PN 23-MAR-2000.
XX 08-SEP-1999; 99WO-US20594.
PF 10-SEP-1998; 98US-0099803.
XX 10-SEP-1998; 98WO-US18824.
PR (GETH) GENENTECH INC.
PA Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI, Botstein D;
XX WPI; 2000-271386/23.
PI P-PSDB; AAY88573.
DR
DR
XX
PT New isolated antibodies which bind to specific polypeptides used for
PT diagnosis and treatment of neoplastic cell growth and proliferation -
XX
PS Example 7; Fig 13; 200pp; English.
XX
CC This sequence represents a human PRO261 nucleotide sequence. PRO261 is a
CC growth factor. The invention relates to isolated antibodies which bind to
CC a polypeptide. The "PRO" polypeptides are encoded by genes which are over
CC expressed in the genome of tumour cells. Vectors and host cells
CC comprising the nucleic acid encoding the antibodies are used in the
CC production of the antibodies. The antibodies and nucleic acids encoding
CC them are used for diagnosing a tumour in a mammal. The antibodies are
CC used for inhibiting the growth of tumour cells and identifying compounds
CC that inhibit a biological or immunological activity of and/or expression
CC of a PRO187, PRO533, PRO214, PRO240, PRO211, PRO230, PRO261, PRO246 or
CC PRO317 polypeptide. The antibody can be used in antibody dependent enzyme
CC mediated prodrug therapy (ADEPT) by conjugating the antibody to a
CC prodrug-activating enzyme which converts a prodrug to an anti-cancer
CC drug. The antibodies can be fluorescently labelled and monitored by light
CC microscopy, flow cytometry or fluorimetry for diagnosis and prognosis of
CC tumours.
XX
SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match
Best Local Similarity 49.2%; Score 434.6; DB 21; Length 1266;
Matches 512; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGCCGAGCCACTGATCCATCTTCTGGCCACTTCTCCTCCTCTC 302
DB 4 GGGGACATGAGAGGACACCCGAGACCCACCTCCTGGCCTTCTCCTCCTCTCCTCCTC 63

QY 303 TCAATGGTGTGTGCCCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACCCACG 362
DB 64 TCAAAGGTGCGTACCCAGCTGTGCCCCGACACCATGTACCTGCCCTGGCCACCTCCCGA 123

QY 363 TGCCCCACAGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTGCTGTAAGTGTGTCACGG 422
DB 124 TGCCCGCTGGAGTACCCCTGGTGTGATGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTG 183

QY 423 AGGCTGGGGAGTCCCTGCGACCACTGCGATGTCTGCGACCCAGCCAGGGCCTGGTTGT 482
DB 184 CGGCTGGGGAGGCCCTGCGACCACTGCGATGTCTGCGACCCAGCCAGGGCCTGGTCTGC 243

QY 483 CAGCCTGGGGAGGCCCTGCGGCGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGT 542
DB 244 CAGCCCGGGGAGGACCCGGTGGCCGGGGCCCTGTGCTCTTGGCAGAGGACGACAGC 303

QY 543 AGCTGTGAGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
DB 304 AGCTGTGAGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGC 363

QY 603 GTCCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTCAGTGAGGATGTG 562

Db 364 ATCCGCTGCCGCTGCGAGGACGGGGCTTACCTGCGTGGCTGTCAGCGAGGATGTG 423
QY 663 CGGCTGCCCGAGCTGGGACTGCCCCACGCCCAAGAGATAACAGGTGCCAGGAAAGTGTGC 722
Db 424 CGGCTGCCCGAGCTGGGACTGCCCCCACCAGGAGGGTTCGAGGTCTTGGGCAAGTGTGC 483
QY 723 CCGGAGTGGGTATGTGACCAGGGAGTGACACCGGGGATCCAGCGCTCCACGGCGCAAGGA 782
Db 484 CCTGAGTGGGTGCGGGCCCAAGAGGGGACTGGGGACCCAGCCCTTCCAGCCCCAAGGA 543
QY 783 CACCAACTTCTGCCCCCTTGTCTACTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842
Db 544 CCCCAGTTTCTGGCCTTGTCTCTTCCCTGCCCTGGTGTGCTGCTGCTGCTGCTGCTGCTG 603

QY 843 ACAGCCTGGGCGCCCTGCTCAACCCACCTGTGGGCTGGGCAT 883
Db 604 ACGGCCTGGGAGCCCTGCTCGACCCACCTGTGGGCTGGGCAT 644

RESULT 10
AAS21403
ID AAS21403 standard; cDNA; 1266 BP.
XX
AC AAS21403;
XX 24-OCT-2001 (first entry)
XX Human cDNA sequence encoding for PRO261 polypeptide.
DE Human secretory and transmembrane; PRO; mammalian; cancer; lung;
XX breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX Homo sapiens.
OS
XX WO200140466-A2.
PN 07-JUN-2001.
XX
XX 01-DEC-2000; 2000WO-US32678.
XX 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.
XX (GETH) GENENTECH INC.
PA
XX

Db 424 CGCTGCCCCAGCTGGGACTGCCCCCACCACAGGAGGGTCGAGGTCTCTGGCAAGTGCTGC 483
QY 723 CCCGAGTGGGTATGTGACCAAGGGAGTGACACCGGGCGGATCCAGCGCTCCACGGCGCAAGGA 782
Db 484 CCTGAGTGGGTGTGCGGCCAAGAGGGGGAGTGGGACCCAGCCCTTCCAGCCCCAAGGA 543
QY 783 CACCAACTTTCTGCCCTTGTCACTCCTGCCTCTGTCTGTGATGCTCTCTGTCCTTCCAAATTTGGAGC 842
Db 544 CCCCAGTTTCTGGCCTTGTCTCTTCCCTGCCCCCTGGTGTCTCCCTGCCCAAGATGGAGC 603
QY 843 ACAGCCTGGGGCCCTGCTCAACCAACCTCTGCGGCTGGGCAT 883
Db 604 ACGGCCTGGGGACCTGCTCGACCACTGTGGGCTGGGCAT 644

RESULT 14
AAV29260
ID AAV29260 standard; DNA; 1285 BP.
XX
AC AAV29260;
XX
DT 14-SEP-1998 (first entry)
XX Human connective tissue growth factor-3 gene.
DE
KW Connective tissue growth factor-3; CTGF-3; human; cancer;
KW arthritis; fibrosis; osteoporosis; diagnosis; therapy; ds.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 9..761
FT /*tag= a
FT sig_peptide 9..65
FT /*tag= b
FT mat_peptide 66..758
FT /*tag= c

XX WO9821236-A1.
PN
XX
PD 22-MAY-1998.
XX
PF 08-NOV-1996; 96WO-US17856.
XX
PR 08-NOV-1996; 96WO-US17856.
XX (HUMA-) HUMAN GENOME SCI INC.

XX Chopra A, Ebner R, Ruben SM;
XX WPI; 1998-297864/26.
DR P-PSDB; AAW37946.
XX

PT Novel human connective tissue growth factor 3 gene - useful for the
PT diagnosis and treatment of e.g. cancer, arthritis, fibrosis,
PT osteoporosis

XX Claim 2; Fig 1A-B; 87pp; English.

XX This nucleotide sequence codes for human connective tissue growth
CC factor-3 (CTGF-3) protein (see AAW37946), a novel member of the
CC growth factor superfamily. It was discovered in a cDNA library
CC derived from human osteoblasts. The gene has also been identified
CC in cDNA libraries from ovary, testis, heart, lung, skeletal muscle,
CC adrenal medulla, adrenal cortex, thymus, prostate, small intestine
CC and colon. A cDNA clone is deposited as ATCC 97756. Also provided
CC are vectors, host cells and recombinant methods for producing
CC CTGF-3 polypeptides. CTGF-3 nucleic acid sequences or their
CC fragments, e.g. primers or probes, can be used to diagnose diseases
CC where CTGF-3 expression is enhanced, e.g. cancer, arthritis,
CC fibrosis or atherosclerosis, or diseases where expression is
CC decreased or increased levels of CTGF-3 can be treated by

CC administering CTGF-3 polypeptides and anti-CTGF-3 antibodies,
CC respectively.
XX
SQ Sequence 1285 BP; 237 A; 418 C; 389 G; 241 T; 0 other;
Query Match 49.2%; Score 434.6; DB 19; Length 1285;
Best Local Similarity 79.9%; Pred. No. 1.6e-110;
Matches 512; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCTC 302
Db 3 GGGGACATGAGAGGCACACCGAAGACCCACCTCTTGGCCTTCTCTCTCTCTCTCTCTC 62
QY 303 TCAATGGTGTGTGCCCAAGTGTGCCGACACCTGTACCTGTCTTGGATGGACACCCAG 362
Db 63 TCAAAGGTGGGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCCGA 122
QY 363 TGCCACACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTTAAAGTGTGTGCACGG 422
Db 123 TGCCCGCTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGCGGCTATGTGCACGG 182
QY 423 AGGCTGGGGAGTCTCTCGACCACTGTCATGTCTCGACCCCAAGAGAGAGTGTGTGTGT 482
Db 183 CGGCTGGGGAGGCTCTCGACCACTGTCACCTGTGCGACGCCAGCGGCTGTGTGTGT 242
QY 483 CAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGT 542
Db 243 CAGCCCCGGGACGAGACCGGTGGCCGGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGC 302
QY 543 AGCTGTGAGGTGATGGCCGAGGTACCTGGATGGAGAGAGACCTTTAAACCAATTGCAGG 602
Db 303 AGCTGTGAGGTGATGGCCGCTGTATCGGGAAGGGAGAGACCTTCCAGCCCCACTGCAGC 362
QY 603 GTCCTGTGCGCTGTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGAGGATGTG 662
Db 363 ATCCGCTGCGCTGTGAGGACGCGGCTTACCTGCGTGCCTGTGAGGAGGATGTG 422
QY 663 CGGCTGCCCAGCTGGGACTGCCACGCCCAAGAGAGATACAGGTGCCAGGAAAGTGTGC 722
Db 423 CGGCTGCCCAGCTGGGACTGCCACGCCCAAGAGGAGGAGGTCGAGGTCTTGGCAAGTGTGC 482
QY 723 CCGGAGTGGGTATGTGACCAAGGAGTGACACCGCGGATCCAGCGTCCACGGCGCAAGGA 782
Db 483 CCTGAGTGGGTGTGCGGCCAAGAGGGGGAGTGGGGACCCAGCCCTTCCAGCCCCAAGGA 542
QY 783 CACCAACTTCTGCCCTTGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 842
Db 543 CCCCAGTTTCTGCGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 602
QY 843 ACAGCCTGGGGCGCCCTGCTCAACCACTGTGGGCTGGGCAT 883
Db 603 ACGGCCTGGGGAGCCCTGCTCGACCACTGTGGGCTGGGCAT 643

RESULT 15
AAV76486
ID AAX76486 standard; DNA; 1293 BP.
XX
AC AAX76486;
XX
DT 06-AUG-1999 (first entry)
XX
DE Human WISP-2 protein nucleotide sequence SEQ ID NO:13.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Homo sapiens.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:55:06 ; Search time 48.04 Seconds
(without alignments)
5636.882 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_883
Perfect score: 883
Sequence: 1 GACGCTCTGATCTCCAGAG.....ACCACCTGTGGGTGGGCAT 883

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	727.6	82.4	1734	4	US-09-182-145-17	Sequence 17, Appl
c 2	727.6	82.4	1734	4	US-09-182-145-18	Sequence 18, Appl
3	434.6	49.2	1293	4	US-09-182-145-13	Sequence 13, Appl
c 4	434.6	49.2	1293	4	US-09-182-145-14	Sequence 14, Appl
5	423	47.9	738	4	US-09-182-145-38	Sequence 38, Appl
6	420.6	47.6	841	4	US-09-182-145-39	Sequence 39, Appl
7	128.4	14.5	2075	1	US-08-167-628-1	Sequence 1, Appli
8	128.4	14.5	2075	1	US-08-386-680-1	Sequence 1, Appli
9	128.4	14.5	2075	1	US-08-459-717-1	Sequence 1, Appli
10	128.4	14.5	2075	1	US-08-712-302-1	Sequence 1, Appli
11	128.4	14.5	2075	2	US-08-880-031-1	Sequence 1, Appli
12	128.4	14.5	2075	3	US-09-097-179-1	Sequence 1, Appli
13	128.4	14.5	2075	4	US-09-080-715-1	Sequence 1, Appli
14	128.4	14.5	2075	4	US-09-142-569-7	Sequence 7, Appli
15	128.4	14.5	2075	5	PCT-US96-08140-1	Sequence 1, Appli
16	128.4	14.5	2998	3	US-09-054-368-1	Sequence 1, Appli
17	128.4	14.5	2998	3	US-09-054-274-1	Sequence 1, Appli
18	128.4	14.5	2998	4	US-09-056-704-1	Sequence 1, Appli
19	125.8	14.2	2350	4	US-09-187-478-1	Sequence 1, Appli
20	124.2	14.1	2350	4	US-09-292-036-1	Sequence 1, Appli
21	123.4	14.0	1418	4	US-09-142-569-3	Sequence 3, Appli
22	120.8	13.7	2267	4	US-09-142-569-5	Sequence 5, Appli
23	117	13.3	1480	4	US-09-142-569-1	Sequence 1, Appli
24	105.6	12.0	1766	4	US-09-182-145-9	Sequence 9, Appli
c 25	105.6	12.0	1766	4	US-09-182-145-10	Sequence 10, Appli
26	102.8	11.6	1128	2	US-08-459-101A-1	Sequence 1, Appli
27	97	11.0	1062	4	US-09-253-316-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-182-145-17
; Sequence 17, Application US/09182145B
; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P1176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 17

; LENGTH: 1734

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-182-145-17

Query Match

Best Local Similarity 82.4%; Score 727.6; DB 4; Length 1734;

Matches 816; Conservatve 91.9%; Pred. No. 1.2e-189;

Mismatches 0; Indels 59; Gaps 4;

QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCCGC 62

Db 13 CGCTCCTGATCTCCAGAGGACCCCGGGCTGGACAGGGCCCTTGGCGAGGCTGCAGCTGC 72

QY 63 TG-GGCAGTGGCTTGGAAATGGAGGCTCTTTATTACTGGGAAGTGAAGAGGCTCC 121

Db 73 TGTGGCAGTAGCTTGGGATGGAGGCTCTTTCTGTGGGAACGTAGGAGGCTGAGAGGCTCC 132

QY 122 TGTACAG---CTTGTCTCTAAAGTCTTAGCAGCTGTGGTGGCTTGCACACTGTGCA 178

Db 133 TGTACAGGCTCCTCTCTAAACTCTTGGCAGCTTGGGTGGCTTGGGCTTGCACACTGTGCA 192

QY 179 GACACCTTCGTGTGGTGGCCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238

Db 193 GACACCTTCTTGGTGGCCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 246

QY 239 ACACGGTGACATGAGGGGACGCCCACTGATCCATCTTCTGGCCACTTCCTTCTCTGCCT 298

Db 247 ACACGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCATTCTCTCTGCAT 306

QY 299 TCTCTCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTAAAGTGTGTGC 358

Db 307 TCTCTCAATGGTGTATTCCACGCTGTGCCAGCACCCCTGTGCCTGTCTTGGACACCACC 366

QY 359 CCAGTGCCACACAGGGGGTACCCCTGTGTGCTGGATGGCTGTGGCTGTAAAGTGTGTGC 418

Db 367 CCAGTGCCACACGGGGGTACCCCTGTGTGCTGGATGGCTGTGGCTGTCTCGAGTGTGTGC 426

QY 419 ACGGAGGCTGGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCCAAGGGCCCTGGT 478

Db 427 ACGGAGGCTGGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCCAAGGGCCCTGGT 486

QY 479 TTGTACGCTGGGGGAGGGCCCTGGCGGCCCATGGGGCTGTGTCTCTTGGATGAGGATGA 538

Db 487 TTGTACGCTGGGGGAGGGCCCAAGTGGCCGTGGTGTGTCTCTTGAAGGATGA 546

QY 539 CGGTAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCAATTG 598

Db 547 CGGGAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGGGAGACCTTTAAACCCAATTG 606

QY 599 CAGGTTCTGTGCCGCTGTGATGACGGTGGCTTCACCTGCCCTGCCCTGTGCAGTGAGGA 658

Db 607 CAGGTTTGTGCCGCTGTGATGACGGTGGTTTCACCTGCCCTGCCCTGTGCAGTGAGGA 666

QY 659 TGTGCGGCTGCCAGCTGGGACTGCCACAGCCGCCCAAGAGAAATACAGGTGCCAGGAAAGT 718

Db 667 TGTGCGGCTGCCAGCTGGGACTGCCACAGCCGCCCAAGAGAAATACAGGTGCCAGGAAAGT 726

QY 719 CTGCCCGAGTGGGTATGTGACCAAGGAGTGA--CACCAGCGGATCCAGCGCTCCACGGC 775

Db 727 CTGCCCGAGTGGGTATGTGACCAAGGAGTGA--CACCAGCGGATCCAGCGCTCCACGGC 786

QY 776 GCAAGGACACCAACTTTCTGCCCTTGTCACTCCTGCCTCTGCTGATGCTCTTGTCCAAA 835

Db 787 CCAAGGACACCAACTTTCTGCCCTTGTCACTCCTGCATCTGCCGATGGCCCTGTCCAAA 846

QY 836 TTGGAGCACAGCCTGGGGCCCTTGCTCAACCACCTGTGGGCTGGGCAT 883

Db 847 CTGGAGCACAGCCTGGGGCCCTTGCTCAACCACCTGTGGGTTGGGCAT 894

RESULT 2

US-09-182-145-18/c

; Sequence 18, Application US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: pl176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 18

; LENGTH: 1734

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-182-145-18

Query Match 82.4%; Score 727.6; DB 4; Length 1734;

Best Local Similarity 91.9%; Pred. No. 1.2e-189;

Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;

QY 3 CGCTTCTGATCTCGAGAGGACCCCTGGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGC 62

Db 1722 CGCTCCTGATCTCGAGAGGACCCCGGGCTGGGACAGGGGCCCTTGGCGAGGCTGCAGCTGC 1663

QY 63 TG-GGCAGTGGCTTGGAAATGGAGGCTCTTTATTACTGGGAACTGAGGAGCTAAGAGGCTCC 121

Db 1662 TGTGGCAGTAGCTGGGATGGAGGCTCTTTCTTGTCTGGAACTGAGGAGCTGAGAGGCTCC 1603

QY 122 TGTGAG---CTTGTCTCTAAAGTCTTAGCAGCTTGTGGTGGCTTGGGCTTCACACACTGTCA 178

Db 1602 TGTACGGCTCTGTCTCTAAACTCTTGGCACTTGGGTGGCTTGGGCTTCACACACTGTCA 1543

QY 179 GACACCTTCTGTGGTGGCCCTCCACGGCCCTCACCTTACAGTTTGAAGCTGGCTCCACAAGGG 238

Db 1542 GACACCTTCTGTGGTGGCCCTCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 1489

QY 239 ACACGGTGACATGAGGGGACGCCCACTGATCCATCTTCTGGCCACTTCCTTCTCTGCCT 298

Db 1488 ACACGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCATTTCTCTCTGCAT 1429

QY 299 TCTCTCAATGGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTCTGGACACCACC 358

Db 1428 TCTCTCAATGGTGTATTCCAGCTGTGCCAGCACCCCTGTGCCTGTCTCTTGGACACCACC 1369

QY 359 CCAGTGCCCAACAGGGGGTACCCCTGGTGTGTGGATGGCTGTGGCTGTCTGTAAAGTGTGTGC 418

Db 1368 CCAGTGCCCAACAGGGGGTACCCCTGGTGTGTGGATGGCTGTGGCTGTCTGTGAGTGTGTGC 1309

QY 419 ACGGAGGCTGGGGGAGTCTCTGGACCACTGCATGTCTGCGACCCCAAGGAGGAGGCTGGT 478

Db 1308 ACGGAGGCTGGGGGAGTCTCTGGACCACTGCATGTCTGCGACCCCAAGGAGGAGTGA 1249

QY 479 TTGTACGCTGGGGGAGGCCCCTGGGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGA 538

Db 1248 TTGTACGCTGGGGGAGGCCCCAGTGGCCGTGGTGTGTGTGCCCTCTCTCGAAGAGGATGA 1189

QY 539 CGGTAGCTGTGAGGTGAATGGCCGCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598

Db 1188 CGGAGCTGTGAGGTGAATGGCCGCGAGGTACCTGGATGGGAGACCTTTAAACCCCAATTG 1129

QY 599 CAGGGTCTGTGCCGCTGTGTGATGACGGTGGCTTACCTGCCTGCCCTGTGCAGTGAGGA 658

Db 1128 CAGGGTTTGTGCCGCTGTGTGATGACGGTGGTTTACCTGCCTGCCCTGTGCAGTGAGGA 1069

QY 659 TGTGCGGCTGCCCAAGCTGGGACTGCCCAAGAGAGAAATACAGGTGCCAGGAAAGTG 718

Db 1068 TGTGCGGCTGCCCAAGCTGGGACTGCCCAAGAGAGAAATACAGGTGCCAGGAAAGTG 1009

QY 719 CTGCCCCGAGTGGGTATGTGACCAGGGAGTGA---CACCAGCGGATCCAGCGCTCCACGGC 775

Db 1008 CTGCCCCGAGTGGGTATGTGACCAGGGAGTGTGACAGCGGCAATCCAGCCCTCTCAGC 949

QY 776 GCAAGGACACCAACTTTCTGCCCTTGTCACTCCTGCCTGTGTGATGCTCTTGTCCAAA 835

Db 948 CCAAGGACACCAACTTTCTGCCCTTGTCACTCCTGCATCTGCGGATGGCCCTGTCCAAA 889

QY 836 TTGGAGCACAGCCTGGGGCCCTTGCTCAACCACCTGTGGGCTGGGCAT 883

Db 888 CTGGAGCACAGCCTGGGGCCCTTGCTCAACCACCTGTGGGTTGGGCAT 841

RESULT 3

US-09-182-145-13

Db 126 TGCCCGCTGGGAGTACCCCTGGTGGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG 185
QY 423 AGGCTGGGGGAGTCCCTGCGACCACTGTCATGTCTGCGACCCAGCCAGGCGCTGTTGT 482
Db 186 CGGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGCGCTGTTGTC 245
QY 483 CAGCCTGGGGCAGGCCCTGGCGGCCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT 542
Db 246 CAGCCCGGGCAGGACCCGGTGGCCGGGGCCCTGTGCTCTTGGCAGAGGACGACAGC 305
QY 543 AGCTGTGAGGTGAATGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 306 AGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 365
QY 603 GTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
Db 366 ATCCGCTGCCGCTGCGAGGACGGCGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 425
QY 663 CGGCTGCCAGCTGGGACTGCCACAGCCCCCAAGAGATAACAGGTGCCAGGAAAGTGTGC 722
Db 426 CGGCTGCCAGCTGGGACTGCCACACCCCAAGAGAGGTGGAGGTCTGGGCAAGTGTGC 485
QY 723 CCCGAGTGGGTATGTGACAGGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGGA 782
Db 486 CCGTGTGGGTGTGCGGCCAAGAGGGGGGACTGGGGACCAAGCCCTTCCA--GCCCAAGGA 543
QY 783 CACCAACTTTCTGCCCTTGTCACTCCTGCCCTCTGCTGATGCTCCTTGTCCAAATTGGAGC 842
Db 544 CCGCAGTTTCTGGCCTTGTCTCTTCCCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
QY 843 ACAGCCTGGGGCCCTGCTCAACCACTGTGGGGCTGGGCAAT 883
Db 604 ACGGCTGGGGACCCCTGCTCGACCACTGTGGGGCTGGGCAAT 644

RESULT 7
US-08-167-628-1
; Sequence 1, Application US/08167628
; Patent No. 5408040
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
US-08-167-628-1

Query Match 14.5%; Score 128.4; DB 1; Length 2075;
Best Local Similarity 52.2%; Pred. No. 7.7e-26;
Matches 338; Conservative 0; Mismatches 301; Indels 9; Gaps 2;

QY 242 CGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTGCTTCT 301
Db 135 CGCGCCAGTATGGGCCCCGTCGCGTGCCTTCGTGCTCCTCGCCCTCTGCAGCG 194
QY 302 CTCAATGGTGTGCCCAAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCAACCCCA 361
Db 195 GCCGCGCTGCGCCAGAACTGCAGCGGGCGCTGCCGCTGCCGACGAGCCGCGCGCG 254
QY 362 GTGCCCCACAGGGGTACCCCTGCTGCTGGATGGCTGTGGCTGTGTAAAGTGTGCACG 421
Db 255 CTGCCCCGGGGCGTGAGCCCTGTGCTGGACGGCTGCGGCTGCTGCCGCTCTGCGCAA 314
QY 422 GAGGCTGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGCGCTGTTG 481
Db 315 GAGCTGGGCGAGTGTGCACCGAGCGGACCCCTGCGACCCGACAAAGGCGCTCTCTG 374
QY 482 TCAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGG 541
Db 375 TGACTTGGCTCCCCGGCCAAACCGCAAGATCGGCGTGTG---CACGCCAAAGATGGTGC 431
QY 542 TAGCTGTGAGGTGAATGGCCGAGGTACCTGGATGGAGAGACTTTAAACCCCAATTGCAG 601
Db 432 TCCCTGCATCTTCGGTGTGTACGGTGTACCGGAGAGAGTCTTCCAGAGCAGCTGCAA 491
QY 602 GGTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCTGCCGCTGTGAGTGGATGT 661
Db 492 GTACCAAGTGCAGTGCCTGGACGGGGCGGTGGGCTGCATGCCCTGTGCAGCATGGACGT 551
QY 662 GCGGCTGCCCCAGCTGGGACTGCCACGCCCCCAAGAGATAACAGGTGCCAGGAAAGTGTG 721
Db 552 TCGTGTGCCAGCCCTGACTGCCCTTCCCGAGGAGGTCAAGCTGCCCGGAAATGCTG 611
QY 722 CCCCAGTGGGTATGTGACCAAGGAGTGACACCGCGGATCCAGCGCTCCACGGCGCAAGG 781
Db 612 CGAGGAGTGGGTGTGTGACGAGCCCAAGACCAACCGTGGTGGGCTGCCCTCGCGGC 671
QY 782 ACACCAACTTCTGCCCTTGTACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
Db 672 TTACCGACTGGAAGACACACGTTTGGCCCAAGACCACTATGATTAGAGCAACTGCCCTGGT 731
QY 836 TTGAGCACAGCCTGGGGCCCTGCTCAACCCACCTGTGGGCTGGGCAT 883
Db 732 CCAGACCACAGAGTGGAGCGCCTGTTCCAAGACCTGTGGGATGGGCAT 779

RESULT 8
US-08-386-680-1
; Sequence 1, Application US/08386680
; Patent No. 5585270
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla


```
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/386,680
FILING DATE: 10-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628
FILING DATE:
APPLICATION NUMBER: US/07/752,427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: DB60R32
FEATURE:
NAME/KEY: CDS
LOCATION: 130..1177
US-08-386-680-1
```

```
Query Match 14.5%; Score 128.4; DB 1; Length 2075;
Best Local Similarity 52.2%; Pred. No. 7.7e-26;
Matches 338; Conservative 0; Mismatches 301; Indels 9; Gaps 2;

QY 242 CGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCCTCTGCTTCT 301
Db 135 CGCGCCAGTATGGGCCCCCGTCCGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 194
QY 302 CTCATGTTGTGTGCCAGCTGTGCGCGGACACCCCTGTACCTGTCTTGGACACCCCA 361
Db 195 GCCGGCGTGGCCAGAACTGACGCGGGCGGTGCCGGTGCCCGGACGAGCCGCGCGCG 254
QY 362 GTGCCACAGGGGTACCCCTGCTGCTGATGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTG 421
Db 255 CTGCCGGCGGGCTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 314
QY 422 GAGCTGGGGAGTCTCTGCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
Db 315 GCAGTGGCGGAGCTGTGCACCGAGCGGACCCCTGCGACCCCGCACAAAGGCTCTTCTG 374
QY 482 TCAGCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTG 541
Db 375 TGACTTGGGCTCCCCGGCCACCGCAAGATCGGCGTGTG---CACCGCCAAAGATGGTGC 431
QY 542 TAGCTGTGAGGTGAATGGCCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG 601
Db 432 TCCCTGCATCTTCGGTGGTACGGGTACCGCAGCGGAGAGTCTTCCAGAGCAGCTGCAG 491
QY 602 GGTCTGTGCCGTGTGATGACGGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
Db 492 GTACAGTGCAGCTGCTGGACGGGGCGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551
QY 662 GCGGTGCCCCAGCTGGGACTGCCACGCCCCCAAGAGATAACAGGTGCCAGGAAAGTGTG 721
Db 552 TCGTGTGCCAGCCCTGACTGCCCCCTTCCCGAGGAGGGTCAAGCTGCCCGGGAATGCTG 611
```

```
QY 722 CCCGAGTGGGTATGTGACGAGGAGTGACACCGGCGATCCAGGCTCCACGGCGCAAGG 781
Db 612 CGAGGAGTGGGTGTGTGACGAGCCCAAGACCAACCGTGGTTGGGCTTCCCTCGCGGC 671
QY 782 ACACCAACTTCTGCCCCCTTGTCACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835
Db 672 TTACCGACTGGAAGACACGTTTGGCCCCAGACCAACTATGATTAGAGCCCAACTGCCT 731
QY 836 TTGGAGCACAGCCTGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCAT 883
Db 732 CCAGACCACAGAGTGGAGCGCCTGTTCCAAGACCTGTGGGATGGGCAT 779
```

```
RESULT 9
US-08-459-717-1
; Sequence 1, Application US/08459717
; Patent No. 5770209
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,717
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,427
; FILING DATE: 30-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
US-08-459-717-1
```

```
Query Match 14.5%; Score 128.4; DB 1; Length 2075;
Best Local Similarity 52.2%; Pred. No. 7.7e-26;
Matches 338; Conservative 0; Mismatches 301; Indels 9; Gaps 2;

QY 242 CGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCCTCTGCTTCT 301
Db 135 CGCGCCAGTATGGGCCCCCGTCCGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 194
QY 302 CTCATGTTGTGTGCCAGCTGTGCGCGGACACCCCTGTACCTGTCTTGGACACCCCA 361
```

Db 195 GCCGGCGCTGGCCAGAACTGCAGCGGGCCGTCGCCGGTGCCCGGACGAGCGCGCGCCGCG 254
QY 362 GTGCCCCACAGGGGTACCCCTGGTGTGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACG 421
Db 255 CTGCCCCGGCGGCTGAGCCTCGTGTGGACGGCTGCGGCTGCTGCCGCTGTGCGCCAA 314
QY 422 GAGGCTGGGGAGTCTGTCGACCACTGCTGCTGCGACCCCGACGCGGCTGGTGTG 481
Db 315 GCAGCTGGCGAGCTGTGCACCGGCGGACCCCTGCGACCCCGCACAGGCGCTTCTG 374
QY 482 TCAGCCTGGGCGAGGCGCTGGCGGCCATGGGCGTGTGTCTCTTGGATGAGGATGACGG 541
Db 375 TGACTTCGGCTCCCGGCCAACCGCAAGATCGGCGTGTG--CACCGCCAAAGATGGTGC 431
QY 542 TAGCTGTGAGGTGAATGGCCGACGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG 601
Db 432 TCCCTGCATCTTCGGTGGTACGGTGTACCGGACGCGGAGATCCTTCCAGAGCAGCTGCAA 491
QY 602 GGTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCGCTGTCAGTGAGGATGT 661
Db 492 GTACCAAGTGACGTGCTGGACGGGGCGGTGGGCTGCATGCCCTGTGACGATGGACGT 551
QY 662 GCGGCTGCCAGCTGGGACTGCCACGCCCCAGAGAGATACAGGTGCCAGAAAGTGTG 721
Db 552 TCGTCTGCCAGCCTGACTGCCCTTCCCGAGAGGGTCAAGCTGCCCGGAAATGCTG 611
QY 722 CCCCAGTGGGTATGTGACCAAGGAGTGACACCGCGCGATCCAGCGCTCCACGGCGCAAGG 781
Db 612 CGAGGAGTGGTGTGACGAGGCCCAAGGACCAACCGTGGTGGGCGCTGCCCTCGCGGC 671
QY 782 ACACCAACTTTCGCCCTGTCACTCCTGCCCTGTC-----TGATGCTCCTTGTCCAA 835
Db 672 TTACCGACTGGAAGACACGTTTGGCCCCAGACCCAACTATGATTAGAGCCCACTGCCTGGT 731
QY 836 TTGGAGCACAGCCTGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCAT 883
Db 732 CCAGACCACAGAGTGGAGGCGCTGTTCCAAGACCTGTGGGATGGGCAT 779

RESULT 10
US-08-712-302-1
; Sequence 1, Application US/08712302
; Patent No. 5783187
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,302
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,680
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.

; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
US-08-712-302-1

Query Match 14.5%; Score 128.4; DB 1; Length 2075;
Best Local Similarity 52.2%; Pred. No. 7.7e-26;
Matches 338; Conservative 0; Mismatches 301; Indels 9; Gaps 2;
QY 242 CGGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCCTTCTCTGCCTTCT 301
Db 135 CGCGCCAGTATGGGCCCCGTCCGGCTCGTGGTGTCTCTCGCCCTCTGCAGCGG 194
QY 302 CTCAATGGTGTGTGCCCAAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCA 361
Db 195 GCCGGCGCTCGGCCAGAACTGCAGCGGGCGCTGCCGGTGCCCGACGAGCGCGCGCGG 254
QY 362 GTGCCCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGTAAGTGTGTGCACG 421
Db 255 CTGCCCGCGGGCGGTGAGCCTCGTGTGGACGCGGACCCCTGCGGCTGCTGCCGCTCTCGCCAA 314
QY 422 GAGGCTGGGGAGTCTCTGCGACCACTGCATGCTGTGCGACCCCGACGCGGCGCTGTTG 481
Db 315 GCAGCTGGGCGAGCTGTGCACCGAGCGGACCCCTGCGACCCCGCACAGGCGCTCTTCTG 374
QY 482 TCAGCCTGGGGCAGGCCCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGG 541
Db 375 TGACTTCGGCTCCCGGCCAACCGCAAGATCGGCGTGTG---CACCGCCCAAGATGGTGC 431
QY 542 TAGCTGTGAGGTGAATGGCCGACGAGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG 601
Db 432 TCCCTGCATCTTCGGTGGTACGGTGTACCGGACGCGGAGAGTCTTCCAGAGCAGCTGCAA 491
QY 602 GGTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCGCTGTGAGTGGATGT 661
Db 492 GTACCAAGTGCATGCTGCCCTGGACGGGCGGTGGGCTGCATGCCCTGTGACGATGGACGT 551
QY 662 GCGGCTGCCAGCTGGGACTGCCACGCCCCAGAGAGATAACAGGTGCCAGAAAGTGTG 721
Db 552 TCGTCTGCCAGCCTGACTGCCCTTCCCGAGAGGGTCAAGCTGCCCGGAAATGCTG 611
QY 722 CCCCAGTGGGTATGTGACCAAGGAGTGACACCGCGCGATCCAGCGCTCCACGGCGCAAGG 781
Db 612 CGAGGAGTGGTGTGACGAGCCCCAAGGACCAACCGTGGTGGGCGCTGCCCTCGCGGC 671
QY 782 ACACCAACTTTCGCCCTGTGCTCACTCCTGCCCTGTC-----TGATGCTCCTTGTCCAA 835
Db 672 TTACCGACTGGAAGACACGTTTGGCCCCAGACCCAACTATGATTAGAGCCCACTGCCTGGT 731
QY 836 TTGGAGCACAGCCTGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCAT 883
Db 732 CCAGACCACAGAGTGGAGGCGCTGTTCCAAGACCTGTGGGATGGGCAT 779

RESULT 11
US-08-880-031-1
; Sequence 1, Application US/08880031
; Patent No. 5916756
; GENERAL INFORMATION:

Db

135

CGCCGCCAGTATGGGCCCCCGTCCGGCTCGCCTTCGTGGTCCCTCGGCCCTCTGCAGCGG

194

QY

302

CTCAATGGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACACCCCA

361

Db

195

GCCGGCCGTCGGCCAGAACTGCAGCGGCGCGTGCCGGTGCCCGACGAGCCGGCGCCGG

254

QY

362

GTGCCCACAGGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTAAAGTGTGTGCACG

421

Db

255

CTGCCCCGGCGGTGAGCCTCGTGTGACGGCTGCGGCTGTGCCCGCTCTGCCGCCAA

314

QY

422

GAGGCTGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCCGAGGGCCTGGTTTG

481

Db

315

GCAGCTGGGGAGCTGTGCACCGAGCGCGGACCCCTGCGACCCCGACAAGGGCCTCTTCTG

374

QY

482

TCAGCCTGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGG

541

Db

375

TGACTTCGGCTCCCCGGCCAAACCGCAAGATCGGCGTGTG---CACCGCCAAAGATGGTGC

431

QY

542

TAGCTGTAGGTGAATGGCCGACAGTACCTGGATGGAGAGACCTTAAACCCAATTGCAG

601

Db

432

TCCCTGCATCTTCGGTGGTACGGGTACCGGACGGGAGAGTCCCTCCAGAGCAGCTGCAA

491

QY

602

GGTCTGTGCCGCTGTGATGACGGTGGCTTCACCTGCCCTGCTGCGAGTGAGGATGT

661

Db

492

GTACCAGTGCACGTGCCCTGGACGGGGCTGGGCTGCATGCCCTGTGCAGCATGGACGT

551

QY

662

GCGGCTGCCAGCTGGGACTGCCCCACGCCCCAAGAGAAATACAGGTGCCAGGAAAGTCTG

721

Db

552

TCGTCTGCCAGCCCTGACTGCCCTTCCCGAGGAGGGTCAAGCTGCCCGGAAATGCTG

611

QY

722

CCCCGAGTGGGTATGTGACCAGGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAG

781

Db

612

CGAGGAGTGGGTGTGTGACGAGCCCAAGGACCAACCGTGGTTGGCCTGCCCTCGCGGC

671

QY

782

ACACCAACTTCTGCCCTTGTCACTCCTGCCTCTGC-----TGATGCTCCTTGTCCAA

835

Db

672

TTACCGACTGGAAGACACGTTTGGCCCCAGACCCCAACTATGATTAGAGCCAACTGCCTGGT

731

QY

836

TTGAGCACAGCCTGGGGCCCCCTGCTCAACCAACCTGTGGGCTGGGCAT

883

Db

732

CCAGACCACAGAGTGGAGCGCCTGTTCCAAGACCTGTGGGATGGGCAT

779

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:36:41 ; Search time 217.668 Seconds
(without alignments)
8368.853 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_883
Perfect score: 883
Sequence: 1 GACGCTTCTGATCTCCAGAG.....ACCACCTGTGGGCTGGGCAT 883

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues 2879534
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	883	100.0	1708	15	US-10-010-408-1 Sequence 1, Appli
2	727.6	82.4	1734	15	US-10-112-267-17 Sequence 17, Appl
3	727.6	82.4	1734	15	US-10-112-267-18 Sequence 18, Appl
4	635	71.9	753	15	US-10-010-408-3 Sequence 3, Appli
5	566	64.1	681	15	US-10-010-408-12 Sequence 12, Appl
6	440	49.8	1337	11	US-09-915-582-30 Sequence 30, Appl
7	439.8	49.8	1352	11	US-09-915-582-14 Sequence 14, Appl
8	434.6	49.2	1266	14	US-10-137-866-319 Sequence 319, App
9	434.6	49.2	1266	14	US-10-146-726-319 Sequence 319, App
10	434.6	49.2	1266	14	US-10-146-727-319 Sequence 319, App
11	434.6	49.2	1266	14	US-10-146-788-319 Sequence 319, App
12	434.6	49.2	1266	14	US-10-152-380-319 Sequence 319, App
13	434.6	49.2	1266	14	US-10-153-934-319 Sequence 319, App
14	434.6	49.2	1266	15	US-10-028-072-319 Sequence 319, App
15	434.6	49.2	1266	15	US-10-121-049-319 Sequence 319, App
16	434.6	49.2	1266	15	US-10-123-904-319 Sequence 319, App

17	434.6	49.2	1266	15	US-10-140-470-319 Sequence 319, App
18	434.6	49.2	1266	15	US-10-175-746-319 Sequence 319, App
19	434.6	49.2	1266	15	US-10-176-918-319 Sequence 319, App
20	434.6	49.2	1266	15	US-10-176-921-319 Sequence 319, App
21	434.6	49.2	1266	15	US-10-137-865-319 Sequence 319, App
22	434.6	49.2	1266	15	US-10-140-474-319 Sequence 319, App
23	434.6	49.2	1266	15	US-10-142-431-319 Sequence 319, App
24	434.6	49.2	1266	15	US-10-143-114-319 Sequence 319, App
25	434.6	49.2	1266	15	US-10-140-002-319 Sequence 319, App
26	434.6	49.2	1266	15	US-10-142-419-319 Sequence 319, App
27	434.6	49.2	1266	15	US-10-123-262-319 Sequence 319, App
28	434.6	49.2	1266	15	US-10-142-423-319 Sequence 319, App
29	434.6	49.2	1266	15	US-10-121-050-319 Sequence 319, App
30	434.6	49.2	1266	15	US-10-141-755-319 Sequence 319, App
31	434.6	49.2	1266	15	US-10-143-032-319 Sequence 319, App
32	434.6	49.2	1266	15	US-10-123-108-319 Sequence 319, App
33	434.6	49.2	1266	15	US-10-123-236-319 Sequence 319, App
34	434.6	49.2	1266	15	US-10-123-261-319 Sequence 319, App
35	434.6	49.2	1266	15	US-10-140-921-319 Sequence 319, App
36	434.6	49.2	1266	15	US-10-140-928-319 Sequence 319, App
37	434.6	49.2	1266	15	US-10-121-045-319 Sequence 319, App
38	434.6	49.2	1266	15	US-10-123-292-319 Sequence 319, App
39	434.6	49.2	1266	15	US-10-123-903-319 Sequence 319, App
40	434.6	49.2	1266	15	US-10-124-819-319 Sequence 319, App
41	434.6	49.2	1266	15	US-10-124-822-319 Sequence 319, App
42	434.6	49.2	1266	15	US-10-140-925-319 Sequence 319, App
43	434.6	49.2	1266	15	US-10-160-498-319 Sequence 319, App
44	434.6	49.2	1266	15	US-10-124-824-319 Sequence 319, App
45	434.6	49.2	1266	15	US-10-127-825A-319 Sequence 319, App

ALIGNMENTS

RESULT 1

US-10-010-408-1

; Sequence 1, Application US/10010408
; Publication No. US20020165185A1

; GENERAL INFORMATION:

; APPLICANT: John J. Castelletto, Jr.
; TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CCN-Like Molecules and Uses Therefor

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/010,408

; FILING DATE: 07-Dec-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/044,273

; FILING DATE: March 19, 1998

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Amy E. Mandragouras

; REGISTRATION NUMBER: 36,207

; REFERENCE/DOCKET NUMBER: MBI-004

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)227-7400

; TELEFAX: (617)742-4214

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

```

;
; LENGTH: 1708 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 249..1001
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-010-408-1

Query Match
Best Local Similarity 100.0%; Score 883; DB 15; Length 1708;
Matches 883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGCTTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCC 60
Db 1 GACGGCTTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCC 60

QY 61 GCTGGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAAGTGGAGGCTAAGAGGCTC 120
Db 61 GCTGGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAAGTGGAGGCTAAGAGGCTC 120

QY 121 CTGTGAGCTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCAGA 180
Db 121 CTGTGAGCTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCAGA 180

QY 181 CACCTTCGTGGTGGCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGGAC 240
Db 181 CACCTTCGTGGTGGCCTCCACGGCCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGGAC 240

QY 241 ACGGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCCTCTGCCTTC 300
Db 241 ACGGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCCTCTGCCTTC 300

QY 301 TCTCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCC 360
Db 301 TCTCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCC 360

QY 361 AGTGCCACAGGGGGTACCCCTGGTGGTGGATGGCTGGCTGTGAAAGTGTGTGCAC 420
Db 361 AGTGCCACAGGGGGTACCCCTGGTGGTGGATGGCTGGCTGTGAAAGTGTGTGCAC 420

QY 421 GGAGGCTGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCCGAGGCGCTGGTTT 480
Db 421 GGAGGCTGGGGAGTCTCTGCGACCACTGCATGTCTGCGACCCCGAGGCGCTGGTTT 480

QY 481 GTCAGCCTGGGGCAGGCCCTGGGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACG 540
Db 481 GTCAGCCTGGGGCAGGCCCTGGGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACG 540

QY 541 GTAGCTGTAGGTGAATGGCCGACGTACCTGGATGGAGAGACCTTAAACCCCAATTGCA 600
Db 541 GTAGCTGTAGGTGAATGGCCGACGTACCTGGATGGAGAGACCTTAAACCCCAATTGCA 600

QY 601 GGGTCTGTGCCGCTGTGATGACGTTGGCTTACCTGCCTGCCCTGTGCAGTGAGGATG 660
Db 601 GGGTCTGTGCCGCTGTGATGACGTTGGCTTACCTGCCTGCCCTGTGCAGTGAGGATG 660

QY 661 TGGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTCT 720
Db 661 TGGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTCT 720

QY 721 GCCCCGAGTGGGTATGTGACCAAGGAGTGACACCGGGGATCCAGCGCTCCACGGGCAAG 780
Db 721 GCCCCGAGTGGGTATGTGACCAAGGAGTGACACCGGGGATCCAGCGCTCCACGGGCAAG 780

QY 781 GACACCAACTTTCTGCCCTTGTCACTCCCTGCCTCTGCTGATGCTCTTGTCCAAATTGGA 840
Db 781 GACACCAACTTTCTGCCCTTGTCACTCCCTGCCTCTGCTGATGCTCTTGTCCAAATTGGA 840

QY 841 GCACAGCCTGGGGCCCTGTCTCAACCACTGTGGGCTGGGCAT 883
Db 841 GCACAGCCTGGGGCCCTGTCTCAACCACTGTGGGCTGGGCAT 883

```

```

Db 841 GCACAGCCTGGGGCCCTGTCTCAACCACTGTGGGCTGGGCAT 883

RESULT 2
US-10-112-267-17
; Sequence 17, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-112-267-17

Query Match
Best Local Similarity 82.4%; Score 727.6; DB 15; Length 1734;
Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;

QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCG 62
Db 13 CGCTCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCCCTTGGGAGGCTGCAGCTGC 72

QY 63 TG-GGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAAGTGGAGAGTAAAGGCTCC 121
Db 73 TGTGGCAGTAGCTTGGGATGGAGGCTTTCTTCTTGGGAACTGAGAGGCTGAGAGGCTCC 132

QY 122 TGTGAG---CTTGTCTTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTACACACTGTCA 178
Db 133 TGTGAGGCTCTGTCTTAAACTCTTGGCACTTGGGCTTGGGCTTGGGCTTACACACTGTCA 192

QY 179 GACACCTTCGTGGTGGCTCCACGGCCCTCACCTTACCTTTCAGGTTTGAAGCTGGCTCCACAAGG 238
Db 193 GACACCTTCTTGGTGGCTCTCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGG 246

QY 239 ACACGGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGSCCACTTCTCTCTGCTCCT 298
Db 247 ACACGGTGACATGAGGGGCAACCCCACTGATCCATCTTCTGSCCACTTCTCTCTGCTCCT 306

QY 299 TCTCTCAATGGTGTGTGCCAGCTGTGCCGGGACACCCCTGTACCTTGTCCCTTGGACACC 358
Db 307 TCTCTCAATGGTGTATTCCAGCTGTGCCCGAGCACCCCTGTGCCCTGTCTTGGACACC 366

QY 359 CCAGTGGCCACAGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTGTAAAGTGTGTGC 418
Db 367 CCAGTGGCCACCGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTGTGTGTGTGTGC 426

QY 419 ACGGAGGCTGGGGAGTCTCTGCGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 478
Db 427 ACGGAGGCTGGGGAGTCTCTGCGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 486

```


QY 479 TTGTCAGCCTGGGGCAGGCCCTGGCGGCCCATGGGGCTGTGTCTCTTTGGATGAGGATGA 538
|||||
Db 487 TTGTCAGCCTGGGGCAGGCCCATGAGCGCGTGGTGTCTGTGCCTCTTCTGAAGAGGATGA 546
QY 539 CGGTAGCTGTGAGGTGAATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598
|||
Db 547 CGGGAGCTGTGAGGTGAATGGCGCAGGTACCTGGATGGGAGACCTTTAAACCCCAATTG 606
QY 599 CAGGFCCTGTGCCGCTGTGATGACGGTGGCTTCACCTGCTGCCGCTGTGCAGTGAGGA 658
|||||
Db 607 CAGGGTTTGTGCCGCTGTGATGACGGTGGTTTCACCTGCTGCCGCTGTGCAGTGAGGA 666
QY 659 TGTGCGCTGCCCAGCTGGGACTGCCACGCCCCCAAGAGAATACAGGTGCCAGGAAAGTG 718
|||||
Db 667 TGTGCGCTGCCCAGCTGGGACTGCCACGCCCCCAGGAGATAACAGGTGCCAGGAAGTG 726
QY 719 CTGCCCCGAGTGGGTATGTGACCAGGGAGTGA--CACCGCGATCCAGCGCTCCACGGC 775
|||||
Db 727 CTGCCCCGAGTGGGTGTGTGACCAGGCAGTGATGCAGCGCGCAATCCAGCCCTCCTCAGC 786
QY 776 GCAGGACACCAACTTTCTGCCCTTGTCACTCCTGCCTCTGCTGATGCTCCTTTGTCCAAA 835
|||||
Db 787 CCAGGACACCAACTTTCTGCCCTTGTCACTCCTGCATCTGCCGATGGCCCTGTTCAAA 846
QY 836 TTGAGCACAGCCTGGGGCCCTGCTCAACCACCTGTGGGCTGGGCAT 883
|||||
Db 847 CTGGAGCACAGCCTGGGGCCCTGCTCAACCACCTGTGGTGGGCAT 894

RESULT 3

US-10-112-267-18/c
; Sequence 18, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 18
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-112-267-18

Query Match 82.4%; Score 727.6; DB 15; Length 1734;
Best Local Similarity 91.9%; Pred. No. 6.2e-203;
Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;
QY 3 CGCTTCTGATCTCCAGAGGACCTGGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCGC 62
|||
Db 1722 CGCTCTGATCTCCAGAGGACCCGGGCTGGGACAGGGGCTTGGCAGGCTGCAGCTGC 1663
QY 63 TG-GGCAGTGGCTTGGAAATGGAGTCTTTATTACTGGGAACTGAGGAGCTAAGAGGCTCC 121
|| ||||| ||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||

Db 1662 TGTGGCAGTAGCTTGGGATGGAGGTCTTTCTTGTGGGAACTGAGAGCTGAGAGGGCTCC 1603
QY 122 TGTCAG--CTTGTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTACACACACTGTCA 178
|||||
Db 1602 TGTGAGGCTCCTGTCTCTAAACTCTTGGCACTTGTGGTGGCTTGGGCTTACACACACTGTCA 1543
QY 179 GACACCTTCGTGGTGGCTTCACGGGCTTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238
|||||
Db 1542 GACACCTTCGTGGTGGCTTCCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 1489
QY 239 ACACGGTGACATGAGGGGCAAGCCCACTGATCCATCTCTGCGCCACTTCCTTCTCTGCGCT 298
Db 1488 ACACGGTGACATGAGGGGCAAGCCCACTGATCCATCTCTGCGCCACTTCCTTCTCTGCGCT 1429
QY 299 TCTCTCAATGGTGTGCTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTTGGACACCA 358
Db 1428 TCTCTCAATGGTGTATTCAGAGCTGTGCCAGCACCCCTGTGCCTGTCTCTTGGACACCA 1369
QY 359 CCAGTGCCCAACAGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTGTAAGTGTGTGC 418
Db 1368 CCAGTGCCCAACAGGGGGTACCCCTGGTGGTGGATGGCTGTGCGAGTGTGTGC 1309
QY 419 ACGGAGGCTGGGGAGTCTTCGACCACTGATGATGATGATGATGATGATGATGATGATGATG 478
Db 1308 ACGGAGGCTGGGGAGTCTTCGACCACTGATGATGATGATGATGATGATGATGATGATGATG 1249
QY 479 TTGTCAGCCTGGGGCAGGCCCTGGCGGCCCATGGGGCTGTGTGTCTCTTGGATGAGGATGA 538
Db 1248 TTGTCAGCCTGGGGCAGGCCCATGAGTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1189
QY 539 CGGTAGCTGTGAGGTGAATGGCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTG 598
Db 1188 CGGAGCTGTGAGGTGAATGGCGCGCAGGTACCTGGATGGGAGAGACCTTTAAACCCCAATTG 1129
QY 599 CAGGGTCTGTGCCGCTGTGATGACGGTGGCTTCACCTGCTGCCGCTGTGCAGTGAGGA 658
Db 1128 CAGGGTTTGTGCCGCTGTGATGACGGTGGTTTCACCTGCTGCCGCTGTGCAGTGAGGA 1069
QY 659 TGTGCGGCTGCCCAGCTGGGACTGCCCCACGCCCCCAAGAGAATACAGGTGCCAGGAAAGTG 718
Db 1068 TGTGCGGCTGCCCAGCTGGGACTGCCCCACGCCCCCAGGAGATAACAGGTGCCAGGAAAGTG 1009
QY 719 CTGCCCCGAGTGGGTATGTGACAGGGAGTGA--CACCGCGATCCAGCGCTCCACGGC 775
Db 1008 CTGCCCCGAGTGGGTGTGTGACAGGCAGTGATGCAGCGCAATCCAGCCCTCCTCAGC 949
QY 776 GCAAGGACACCAACTTTCTGCCCTTGTCACTCCTGCTGCCCTGTGCTGATGCTCCTTGTCCAAA 835
Db 948 CCAAGGACACCAACTTTCTGCCCTTGTCACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 889
QY 836 TTGAGCACAGCCTGGGGCCCTGCTCAACCACCTGTGCGGCTGGGCAT 883
Db 888 CTGAGCACAGCCTGGGGCCCTGCTCAACCACCTGTGSGTGGGCAT 841

RESULT 4

US-10-010-408-3
; Sequence 3, Application US/10010408
; Publication No. US20020165185A1
; GENERAL INFORMATION:
; APPLICANT: John J. Castellet, Jr.
; TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules and Uses Therefor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 753 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 1..750
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-010-408-3

Query Match 71.9%; Score 635; DB 15; Length 753;
Best Local Similarity 100.0%; Pred. No. 6.6e-176;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 249 ATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCCTTCTGCTTCTCTCAATG 308
Db 1 ATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCCTTCTGCTTCTCTCAATG 60
QY 309 GTGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCCCA 368
Db 61 GTGTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCCCA 120
QY 369 CAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTTGGATGAGGAGGCTG 428
Db 121 CAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTTGGATGAGGAGGCTG 180
QY 429 GGGAGTCTGCGACCACCTGCATGTCTTGGATGAGGAGGCTGTTGTGAGCCT 488
Db 181 GGGAGTCTGCGACCACCTGCATGTCTTGGATGAGGAGGCTGTTGTGAGCCT 240
QY 489 GGGCAGGCCCTGGCGCCATGGGCTGTGTCTTGGATGAGGAGGCTGTTGTGAGCCT 548
Db 241 GGGCAGGCCCTGGCGCCATGGGCTGTGTCTTGGATGAGGAGGCTGTTGTGAGCCT 300
QY 549 GAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCCTG 608
Db 301 GAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCCTG 360
QY 609 TGCCCGTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGAGTGAAGTGTGCGGCTG 668
Db 361 TGCCCGTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGAGTGAAGTGTGCGGCTG 420
QY 669 CCCAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGCCCGGAG 728
Db 421 CCCAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGCCCGGAG 480
QY 729 TGGGTATGTGACGAGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGGACACCAA 788
Db 481 TGGGTATGTGACGAGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGGACACCAA 540
QY 789 CTTTCTGCCCTTGTCACTCCTGCCTCTGCTGATGCTCCTTGTCCAAATTGGAGCACAGCC 848

Db 541 CTTTCTGCCCTTGTCACTCCTGCCTCTGCTGATGCTCCTTGTCCAAATTGGAGCACAGCC 600
QY 849 TGGGGCCCTGCTCAACCACTGTGGCTGGGCAT 883
Db 601 TGGGGCCCTGCTCAACCACTGTGGCTGGGCAT 635
RESULT 5
US-10-010-408-12
Sequence 12, Application US/10010408
Publication No. US20020165185A1
GENERAL INFORMATION:
APPLICANT: John J. Castellet, Jr.
TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CCN-Like Molecules and Uses Therefor
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 681 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 1..681
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-010-408-12

Query Match 64.1%; Score 566; DB 15; Length 681;
Best Local Similarity 100.0%; Pred. No. 9.9e-156;
Matches 566; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 318 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCCCAAGGGGTA 377
Db 1 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCCCAAGGGGTA 60
QY 378 CCCCTGGTGGATGGCTGTGGCTGTAAAGTGTGCAGGAGGCTGGGGAGTCC 437
Db 61 CCCCTGGTGGATGGCTGTGGCTGTAAAGTGTGCAGGAGGCTGGGGAGTCC 120
QY 438 TCGGACCACCTGCATGTCTGCGACCCCAAGGCGCTGGTTGTTCAGCCTGGGGCAGGC 497
Db 121 TCGGACCACCTGCATGTCTGCGACCCCAAGGCGCTGGTTGTTCAGCCTGGGGCAGGC 180

QY	498	CCTGGCGGCCATGGGCGTGTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT	557
Db	181	CCTGGCGGCCATGGGCGTGTGTGTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT	240
QY	558	GGCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCCGTGTGCCGCTGT	617
Db	241	GGCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCCGTGTGCCGCTGT	300
QY	618	GATGACGGTGGCTTCACCTGCCTGCCGCTGTGCAGTGAGGATGTGCGGGCTGCCACGCTGG	677
Db	301	GATGACGGTGGCTTCACCTGCCTGCCGCTGTGCAGTGAGGATGTGCGGGCTGCCACGCTGG	360
QY	678	GACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTCTGCCCCGAGTGGGTATGT	737
Db	361	GACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTCTGCCCCGAGTGGGTATGT	420
QY	738	GACCAGGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTTGCC	797
Db	421	GACCAGGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTTGCC	480
QY	798	CTTGTCACTCCTGCCCTGTGCTGATGCTCCTTGTCCAAATTGGAGCACAGCCTGGGGCCCC	857
Db	481	CTTGTCACTCCTGCCCTGTGCTGATGCTCCTTGTCCAAATTGGAGCACAGCCTGGGGCCCC	540
QY	858	TGCTCAACCACCTGTGGGCTGGGCAT	883
Db	541	TGCTCAACCACCTGTGGGCTGGGCAT	566

RESULT 6

```

US-09-915-582-30
; Sequence 30, Application US/09915582
; Patent No. US20020120103A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Pro
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1337
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1337)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-915-582-30

```

Db	114	CCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCCGATGCCCGTGGGAG	173
QY	376	TACCCCTGGTGGATGGCTGTGGCTGCTGTAAAGTGTGTCAACGGAGGCTGGGGGAGT	435
Db	174	TACCCCTGGTGGATGGCTGTGGCTGCTGCCGGGTATGTCAACGGCGGCTGGGGGAGC	233
QY	436	CCTGGGACCACTTCGATGTCTGGACCCCGAGCCAGGGCCCTGGTTGTTCAGCCTGGGGCAG	495
Db	234	CCTGGGACCACTTCAGCTCTGGACGCCAGCCAGGGCCCTGGTCTGCCAGCCCCGGGGCAG	293
QY	496	GCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGA	555
Db	294	GACCCGGTGGMCGGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCAGCTGTGAGGTGA	353
QY	556	ATGGCCGCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTCAGAGGTCTGTGCCCGCT	615
Db	354	ACGGCCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCCGCT	413
QY	616	GTGATGACGGTGGCTTCACCTGCCTGCCGCTGTGCAGTGAGGATGTGCGGCTGCCCAGCT	675
Db	414	GCGAGGACGGCGGCTTCACCTGCGTGCCGCTGTGCAGCGAGGATGTGCGGCTGCCCAGCT	473
QY	676	GGGACTGCCCCACGCCCAAGAGAATACAGGTGCCAGGAAAGTCTCTGCCCCGAGTGGGTAT	735
Db	474	GGGACTGCCCCCAACCCAGGAGGGTCGAGGTCTCTGGGCAAGTCTGCCCTGAGTGGGTGT	533
QY	736	GTGACCAAGGAGTGACACCGGCGGATCCAGCGCTCCACGGCGCAAGGACACCACTTTCTG	795
Db	534	GCGGCCAAGGAGGGGACTGGGGACCCAGCCCCCTTCCAGCCCCAAGGACCCCACTTTCTG	593
QY	796	CCCTTGTCACTCCTGCCTCTGCTGATGCTCCTTGTCCAAATTTGGAGCACAGCCTGGGGCC	855
Db	594	GCCTTGTCTCTTCCCTGCCCCCTGGTGTCCCTGTGCCCAAGATGGAGCACGGCCTGGGGAC	653
QY	856	CCTGCTCAACCACTGTGGGCTGGGCAT	883
Db	654	CCTGCTCGACCACTGTGGGCTGGGCAT	681

RESULT 7

```

US-09-915-582-14
; Sequence 14, Application US/09915582
; Patent No. US20020120103A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 14
; LENGTH: 1352
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-915-582-14

```

```

Query Match          49.8%; Score 439.8; DB 11; Length 1352;
Best Local Similarity 78.4%; Pred. No. 9.le-119;
Matches 543; Conservative 0; Mismatches 142; Indels 8; Gaps 1;

QY      191 GTGGCCTCCAGGGCCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGACACGGTGACAT 250
      || || || | ||| ||||| | ||||| ||| ||| ||||| ||| ||| |||||
db       3  GTCGGCTTCACAGTTTCACCTTCAGGCTCAAGCTGGCTCTGCA-----GGGGACAT 54

```



```
QY 251 GAGGGGAGCCCACTGATCCATCTTCTGTGCCACTTCTTCTGCTGCTTCTCTCAATGGT 310
    ||| ||| || || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55 GAGAGGACACCGAAGACCCACCTCTCTGGCCTTCTCCCTCTCTGCTCTCTCAAAGGT 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 311 GTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCCCA 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 GCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCGCT 174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 371 GGGGTACCCCTGGTGGTGGATGGGTGGCTGTAAAGTGTGTGCACGGAGGCTGG 430
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 GGGAGTACCCCTGGTGGTGGATGGGTGGCTGTGGCTGTGCCGGGTATGTGCACGGCGCTGG 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 431 GGAGTCTCGCACCACTGATGTCTCGGACCCAGCCAGCCAGGCTGTGTTCAGCCCTGG 490
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 GGAGCCCTCGCACCACTCCACGTCTCGGACGCCAGCCAGGCGTGGTCTGCCAGCCCG 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 491 GGCAGGCCCTGGCGCCATGGGCTGTGTCTTGGATGAGGATGACGGTAGCTGTGA 550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 GGCAGGCCCTGGTGACGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCAGCTGTGA 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 551 GGTAATGGCCGCGAGGTACCTGGATGGAGAGACCTTTAAACCAATTGCAGGGTCTGTG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 GGTGAAGCGCCGCTGTATCGGAAGGGGAGACCTTCCAGCCCACTGCAGCATCCGCTG 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 611 CCGCTGTGATGACGGTGGCTTCACTGCGCTGCCGTGTGCAGTGAGGATGTGCGGCTGCC 670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 415 CCGCTGGAGGACGGCGCTTCACTGCGTGGCTGTGCAGGAGGATGTGCGGCTGCC 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 671 CAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGCCCGAGTG 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 CAGCTGGGACTGCCCAACCCAGGAGGGTGCAGGTCTGGGCAAGTGTGCCCTGAGTG 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 731 GGTATGTACAGGAGTACACCGGCGGATCCAGCGCTCCAGGCGCAAGGACACCAACT 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 GGTGTGGCCCAAGGAGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGGACCCCAAGT 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 791 TTCTGCCCTGTCACTCCTGCTGCTGTGATGCTCTTGTCAAATGGAGCACAGCCTG 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 TTCGTGCTGTCTCTTCCCTGCCCTGGTGTCCCTGCCAGAAATGGAGCACGGCCTG 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 851 GGGCCCTGTCAACCACTGTGGGCTGGGCAT 883
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 GGGACCTGTCTGACCACTGTGGGCTGGGCAT 687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-10-137-866-319
; Sequence 319, Application US/10137866
; Publication No. US20030129689A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C151
; CURRENT APPLICATION NUMBER: US/10/137,866
; CURRENT FILING DATE: 2002-05-03
```

```
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
```

;	PRIOR FILING DATE:	1997-12-11
;	PRIOR APPLICATION NUMBER:	60/069278
;	PRIOR FILING DATE:	1997-12-11
;	PRIOR APPLICATION NUMBER:	60/069334
;	PRIOR FILING DATE:	1997-12-11
;	PRIOR APPLICATION NUMBER:	60/069694
;	PRIOR FILING DATE:	1997-12-16
;	PRIOR APPLICATION NUMBER:	60/072320
;	PRIOR FILING DATE:	1998-01-23
;	PRIOR APPLICATION NUMBER:	60/073612
;	PRIOR FILING DATE:	1998-02-04
;	PRIOR APPLICATION NUMBER:	60/074086
;	PRIOR FILING DATE:	1998-02-09
;	PRIOR APPLICATION NUMBER:	60/074092
;	PRIOR FILING DATE:	1998-02-09
;	PRIOR APPLICATION NUMBER:	60/077791
;	PRIOR FILING DATE:	1998-03-12
;	PRIOR APPLICATION NUMBER:	60/078910
;	PRIOR FILING DATE:	1998-03-20
;	PRIOR APPLICATION NUMBER:	60/079294
;	PRIOR FILING DATE:	1998-03-25
;	PRIOR APPLICATION NUMBER:	60/079663
;	PRIOR FILING DATE:	1998-02-27
;	PRIOR APPLICATION NUMBER:	60/079728
;	PRIOR FILING DATE:	1998-03-27
;	PRIOR APPLICATION NUMBER:	60/080165
;	PRIOR FILING DATE:	1998-03-31
;	PRIOR APPLICATION NUMBER:	60/081203
;	PRIOR FILING DATE:	1998-04-09
;	PRIOR APPLICATION NUMBER:	60/081229
;	PRIOR FILING DATE:	1998-04-09
;	PRIOR APPLICATION NUMBER:	60/081695
;	PRIOR FILING DATE:	1998-04-14
;	PRIOR APPLICATION NUMBER:	60/081817
;	PRIOR FILING DATE:	1998-04-15
;	PRIOR APPLICATION NUMBER:	60/081818
;	PRIOR FILING DATE:	1998-04-15
;	PRIOR APPLICATION NUMBER:	60/082999
;	PRIOR FILING DATE:	1998-04-24
;	PRIOR APPLICATION NUMBER:	60/083322
;	PRIOR FILING DATE:	1998-04-28
;	PRIOR APPLICATION NUMBER:	60/083545
;	PRIOR FILING DATE:	1998-04-29
;	PRIOR APPLICATION NUMBER:	60/084600
;	PRIOR FILING DATE:	1998-05-07
;	PRIOR APPLICATION NUMBER:	60/084627
;	PRIOR FILING DATE:	1998-05-07
;	PRIOR APPLICATION NUMBER:	60/084637
;	PRIOR FILING DATE:	1998-05-07
;	PRIOR APPLICATION NUMBER:	60/085149
;	PRIOR FILING DATE:	1998-05-12
;	PRIOR APPLICATION NUMBER:	60/085323
;	PRIOR FILING DATE:	1998-05-13
;	PRIOR APPLICATION NUMBER:	60/085338
;	PRIOR FILING DATE:	1998-05-13
;	PRIOR APPLICATION NUMBER:	60/085339
;	PRIOR FILING DATE:	1998-05-13
;	PRIOR APPLICATION NUMBER:	60/085579
;	PRIOR FILING DATE:	1998-05-15
;	PRIOR APPLICATION NUMBER:	60/085697
;	PRIOR FILING DATE:	1998-05-15
;	PRIOR APPLICATION NUMBER:	60/085704
;	PRIOR FILING DATE:	1998-05-15
;	PRIOR APPLICATION NUMBER:	60/086414
;	PRIOR FILING DATE:	1998-05-22
;	PRIOR APPLICATION NUMBER:	60/086430
;	PRIOR FILING DATE:	1998-05-22
;	PRIOR APPLICATION NUMBER:	60/087106
;	PRIOR FILING DATE:	1998-05-28
;	PRIOR APPLICATION NUMBER:	60/088026
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088730
;	PRIOR FILING DATE:	1998-06-10

;	PRIOR APPLICATION NUMBER:	60/088741
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088810
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088858
;	PRIOR FILING DATE:	19/98-06-11
;	PRIOR APPLICATION NUMBER:	60/089532
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089599
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089907
;	PRIOR FILING DATE:	1998-06-18
;	PRIOR APPLICATION NUMBER:	60/089947
;	PRIOR FILING DATE:	1998-06-19
;	PRIOR APPLICATION NUMBER:	60/090349
;	PRIOR FILING DATE:	1998-06-23
;	PRIOR APPLICATION NUMBER:	60/090429
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090445
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090538
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090863
;	PRIOR FILING DATE:	1998-06-26
;	PRIOR APPLICATION NUMBER:	60/091360
;	PRIOR FILING DATE:	1998-07-01
;	PRIOR APPLICATION NUMBER:	60/091519
;	PRIOR FILING DATE:	1998-07-02
;	PRIOR APPLICATION NUMBER:	60/091982

Query Match 49.2%; Score 434.6; DB 14; Length 1266;
Best Local Similarity 79.9%; Pred. NO. 3e-117;
Matches 512; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY	243	GGTGACATGAGGGCAGCCACTGATCCATCTTCTGGCCACTTCCTCCCTCTGCCTTCTC	302
DB	4	GGGGACATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTC	63
QY	303	TCAATGGTGTGTGCCACAGCTGTGCCGGACACCCCTGTACCTGTCTCTGGACACCAACCCAG	362
DB	64	TCAAAGTGGCTACCCAGCTGTGCCCCACACCATGTACCTGCCCCCTGGCCACCTCCCCGA	123
QY	363	TGCCCCAAGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTGCTGTAAAGTGTGTGCACGG	422
DB	124	TGCCCCGTGGAGTACCCCTGGTGGTGGATGGCTGTGGCTGTGCTGGCGGGTATGTGCACGG	183
QY	423	AGGCTGGGGAGTCTTGGACCACTGCATGTCTGCGACCCAGCCAGGGCCCTGGTTTGT	482
DB	184	CGGCTGGGGAGCCCTGGACCAACTCCACGTCTGCGACGCCAGCAGGGCCTGGTCTGC	243
QY	483	CAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT	542
DB	244	CAGCCCGGGCAGGACCCGGTGGCCGGGGGGCCCTGTGCCCTCTTGCAGAGGACGACAGC	303
QY	543	AGCTGTGAGGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG	602
DB	304	AGCTGTGAGGTGAACGGCGCCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC	363
QY	603	GTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCGCTGTGCAGTGAGGATGTG	662
DB	364	ATCCGCTGCCGCTGCGAGGACGGCGGCTTACCTGCGTGCCTGTGTCAGCGAGGATGTG	423
QY	663	CGGCTGCCCCAGCTGGGACTGCCACGCCCCCAAGAGAATACAGGTGCCAGGAAAAGTGTCTGC	722
DB	424	CGGCTGCCCCAGCTGGGACTGCCCCCACCACCCAGGAGGGTCTGAGGTCTCTGGGCAAGTGTCTGC	483
QY	723	CCCCAGTGGGTATGTGACCAGGGAGTGACACCCGGCGATCCAGCGCTCCACGGCGCAAGGA	782
DB	484	CCTGAGTGGGTGTGGGGCCCAAGGAGGGGGACTGGGGACCCAGCCCTTCCAGCCCCAAGGA	543
QY	783	CACCAACTTTCTGCCCTTGTCACTCCTGCCTCTGCTGTATGCTCCTTGTCCAAATTGGAGC	842
DB	544	CCCCAGTTTTCTGGCCTGTCTCTTCCCTGCCCCCTGGTGTCCCTTGGCCAGAAATGGAGC	603


```
Db      244 CAGCCCGGGCAGGACCCGGTGGCCGGGGGCCCTGTGCCCTTTGGCAGAGGACGACAGC 303
QY      543 AGCTGTGAGTGAATGGCCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db      304 AGCTGTGAGTGAACGGCCGCTGTATCGGAAGGGGAGACCTTCCAGCCCCCACTGCAGC 363
QY      603 GTCTGTGCGCTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGCAGTGAGGATGTG 662
Db      364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTACCTGCGTGGCTGTGCAGCGAGGATGTG 423
QY      663 CGGCTGCCAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGC 722
Db      424 CGGCTGCCAGCTGGGACTGCCCCCAAGAGGGTGCAGGGTGCAGGGCAAGTGTGC 483
QY      723 CCCGAGTGGGTATGTGACAGGGAGTGACACCGGCGCATCCAGCGTCCCAAATTGGAGC 782
Db      484 CCTGAGTGGGTGTCGGCCCAAGAGGGGACTGGGGACCCAGCCCTTCCAGCCCCAAGGA 543
QY      783 CACCAACTTCTGCCCTTGTGTCACCTGCTGCTGCTGATGCTCCTTGTCCAAATTGGAGC 842
Db      544 CCCCAGTTTCTGGCCTTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCCAAGATGGAGC 603
QY      843 ACAGCCTGGGCCCCCTGCTCAACCACTGTGGGCTGGGCAT 883
Db      604 ACGGCCTGGGACCCCTGCTCGACCACTGTGGGCTGGGCAT 644
```

RESULT 11

```
US-10-146-788-319
; Sequence 319, Application US/10146788
; Publication No. US20030129693A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C322
; CURRENT APPLICATION NUMBER: US/10/146,788
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-788-319
```

```
Query Match      49.2%; Score 434.6; DB 14; Length 1266;
Best Local Similarity 79.9%; Pred. No. 3e-117;
Matches 512; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY      243 GGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCCTTCTGCTCTC 302
Db      4 GGGGACATGAGAGGCGACACCGAAGACCCACCTCTGCGCTTCTCCCTCTGCTCCTC 63
QY      303 TCAATGFGTGTGCCCGACAGCTGTGCGGACACACCTGTACCTGTCTGAGACACCCAG 362
```

```
Db      64 TCAAAGGTGCGTACCCAGCTGTGCCCCGACACACCATGTACCTGCCCTTGGCCACCTCCCCGA 123
QY      363 TGCCACAGGGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGG 422
Db      124 TGCCCGCTGGGAGTACCCCTGGTGGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG 183
QY      423 AGGCTGGGGAGTCTCTGCGACCACTGCTGATGTCTGCGACCCCGAGCGGGCCTGGTTGT 482
Db      184 CGGCTGGGGAGCCCTGCGACCACTGCTGACGTCTGCGACCGCAGCGAGGGCCTGGTCTGC 243
QY      483 CAGCCTGGGCGAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGT 542
Db      244 CAGCCCGGGCGAGGACCCCGGTGGCCGGGGGCCCTGTGCTCTTGGCAGAGGACGACAGC 303
QY      543 AGCTGTGAGTGAATGGCCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db      304 AGCTGTGAGTGAACGGCCGCTGTATCGGAAGGGGAGACCTTCCAGCCCCCAAGGA 363
QY      603 GTCCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCTGCCCTGTGCGAGTGCAGTGCAGTGTG 662
Db      364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTACCTGCGTGGCTGCGAGCGAGGATGTG 423
QY      663 CGGCTGCCAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGC 722
Db      424 CGGCTGCCAGCTGGGACTGCCCCCAAGAGGGGACTGGGGACCCCTTCCAGCCCCAAGGA 483
QY      723 CCCGAGTGGGTATGTGACAGGGAGTGACACCGGCGCATCCAGCGTCCACGGCGCAAGGA 782
Db      484 CCTGAGTGGGTGTCGGCCCAAGAGGGGACTGGGGACCCAGCCCTTCCAGCCCCAAGGA 543
QY      783 CACCAACTTCTGCCCTTGTGTCACCTCCTGCCCTCTGCTGATGCTCCTTGTCCAAATTGGAGC 842
Db      544 CCCCAGTTTCTGGCCTTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCCAAGATGGAGC 603
QY      843 ACAGCCTGGGCCCCCTGCTCAACCACTGTGGGCTGGGCAT 883
Db      604 ACGGCCTGGGACCCCTGCTCGACCACTGTGGGCTGGGCAT 644
```

RESULT 12

```
US-10-152-380-319
; Sequence 319, Application US/10152380
; Publication No. US20030129694A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C397
; CURRENT APPLICATION NUMBER: US/10/152,380
; CURRENT FILING DATE: 2002-05-21
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-152-380-319
```


;
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809

;
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15

Db	184	CGGCTGGGGAGCCCTGCGACCAACTCCACGCTCGCGACGCCAGCCAGGGCCTGGTCTGC	243
QY	483	CAGCCTGGGGCAGGCCCTGGCGGCCCATGGGGCTGTGTCTCTCTTGGATGAGGATGACGGT	542
Db	244	CAGCCCGGGCAGGACCCGGTGGCCGGGGGCCCTGTGCCCTCTTGGCAGAGGACGACAGC	303
QY	543	AGCTGTGAGGTGAATGCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGG	602
Db	304	AGCTGTGAGGTGAACGSCCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCCACTGCAGC	363
QY	603	GTCTGTGCCGCTGTGATGACGGTGGCTTCACCTGCCCTGCCGCTGTGCAGTGAGGATGTG	662
Db	364	ATCCGCTGCCGCTCGAGGACGGCGGCTTCACCTGCCGTGCCGCTGTGCAGCGAGGATGTG	423
QY	663	CGGCTGCCCCAGCTGGGACTGCCCCACGCCCCCAAGAGAATACAGGTGCCAGGAAAGTGTCTGC	722
Db	424	CGGCTGCCAGCTGGGACTGCCCCCAACCCACAGGAGGGTCTGAGGTCTGGGCAAGTGTCTGC	483
QY	723	CCCGAGTGGGTATGTGACCAAGGGAGTGACACCGCGGATCCAGCGCTCCACGGCGCAAGGA	782
Db	484	CCTGAGTGGGTGTGCGSCCAAGGAGGGGGGACTGGGGACCCAGCCCCCTTCCAGCCCCAAGGA	543
QY	783	CACCAACTTCTGCCCTTGTTCACCTCCTGCCCTCTGCTGATGCTCCTTGTCCAAATTGGAGC	842
Db	544	CCCCAGTTTTCTGGCCCTGTCTCTTCCCTGCCCCCTGGTGTCCCCCTGCCCAGAATGGAGC	603
QY	843	ACAGCCTGGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCAT	883
Db	604	ACGGCCTGGGGACCCCTGCTCGACCACTGTGGGCTGGGCAT	644

Search completed: July 28, 2003, 21:31:53
Job time : 218.668 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 18:02:17 ; Search time 1565.76 seconds
(without alignments)
9133.322 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_883
Perfect score: 883
Sequence: 1 GACGCTTCTGATCTCCAGAG.....ACCACCTGTGGGCTGGGCAT 883

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	429.2	48.6	1058	BM805088	BM805088 AGENCOURT
2	425.8	48.2	940	BQ937887	BQ937887 AGENCOURT
3	419.8	47.5	979	BQ279131	BQ279131 AGENCOURT
4	413	46.8	1073	BQ073722	BQ073722 AGENCOURT
5	400.6	45.4	750	BM043988	BM043988 603620978
6	400.6	45.4	1006	BM921531	BM921531 AGENCOURT

7	394.6	44.7	886	13	BI822142	BI822142 603039845
8	391.8	44.4	651	13	BI457141	BI457141 603185392
9	385.2	43.6	537	14	BQ560868	BQ560868 H4067A01-
10	374.6	42.4	928	13	BI161474	BI161474 602864871
11	367.2	41.6	800	13	BI826781	BI826781 603077268
12	367.2	41.6	888	13	BI825652	BI825652 603072631
13	366	41.4	620	13	BG928868	BG928868 HNC57-1-D
14	363.4	41.2	916	13	BI457367	BI457367 603185689
15	361.2	40.9	1166	13	BM543799	BM543799 AGENCOURT
16	349.6	39.6	749	9	AL555144	AL555144 603185689
17	336.4	38.1	1251	14	BQ961357	BQ961357 AGENCOURT
18	335.8	38.0	424	10	BB849097	BB849097 603185392
19	307	34.8	790	13	BM046275	BM046275 603626068
20	306	34.7	1022	14	BQ952960	BQ952960 AGENCOURT
21	304.6	34.5	452	10	BE481184	BE481184 166529 BA
22	297.8	33.7	933	14	BQ278961	BQ278961 AGENCOURT
23	290.6	27.2	489	14	BM751866	BM751866 K-EST0028
24	226.4	25.6	618	12	BG538695	BG538695 602566932
25	223.6	25.3	792	13	BI823598	BI823598 603040962
26	223.2	25.3	405	12	BG900069	BG900069 HOA51-1-A
27	218	24.7	657	13	BM488499	BM488499 pgm2n.pk0
28	209.6	23.7	380	12	BG900020	BG900020 HOA48-1-G
29	128.6	14.6	936	14	BQ880159	BQ880159 AGENCOURT
30	128.2	14.5	887	14	BQ894536	BQ894536 AGENCOURT
31	127.6	14.5	1085	13	BM558306	BM558306 AGENCOURT
32	126.8	14.4	1117	12	BF982720	BF982720 602304922
33	126.6	14.3	828	9	AL543019	AL543019 602304922
34	125.6	14.2	940	14	BQ706630	BQ706630 AGENCOURT
35	125.4	14.2	811	9	AU133455	AU133455 602318826
36	125	14.2	505	12	BF037009	BF037009 601456766
37	125	14.2	826	12	BG116760	BG116760 602318826
38	125	14.2	877	9	AL546864	AL546864 602318826
39	125	14.2	888	9	AL544380	AL544380 602318826
40	125	14.2	896	14	BQ719732	BQ719732 AGENCOURT
41	125	14.2	942	9	AL544453	AL544453 602318826
42	125	14.2	943	9	AL553324	AL553324 602318826
43	125	14.2	982	9	AL550752	AL550752 602318826
44	125	14.2	1004	13	BM472832	BM472832 AGENCOURT
45	125	14.2	1096	13	BM554782	BM554782 AGENCOURT

ALIGNMENTS

RESULT 1
BM805088
LOCUS BM805088 1058 bp mRNA linear EST 05-MAR-2002
DEFINITION AGENCOURT_6490429 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5725937
5', mRNA sequence.
ACCESSION BM805088
VERSION BM805088.1 GI:19121911
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1058)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12717 row: j column: 18
High quality sequence stop: 697.
FEATURES
source Location/Qualifiers
1..1058

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5725937"
/clone_lib="NIH_MGC_125"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT 139 a 423 c 294 g 200 t 2 others
ORIGIN
Query Match 48.6%; Score 429.2; DB 14; Length 1058;
Best Local Similarity 79.3%; Pred. No. 3.7e-99;
Matches 509; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 242 CGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCCCTGCTCTTCT 301
Db 2 CAGGGGAATGAGAGGACACCGAAGACCCACCTCTGCTTCTCTCCCTCTCTGCTCTCT 61
QY 302 CTCAATGGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTTGGACACCCCA 361
Db 62 CTCAAGGTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCG 121
QY 362 GTGCCACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 421
Db 122 ATGCCCGTGGGAGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 181
QY 422 GAGGCTGGGGAGTCTCGACACCACTGCATGTCTGCGACCCAGCCAGGCGCTGGTTG 481
Db 182 GCGGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGCGCTGGTCTG 241
QY 482 TCAGCCTGGGCGAGCCCTGGCGCCATGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 541
Db 242 CCAGCCCGGGCAGGACCCGGTGGACGGGGGCGCTGTGCTCTTGGCAGAGACGACAG 301
QY 542 TAGCTGTGAGGTGATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATTGCAG 601
Db 302 CAGCTGTGAGGTGAACGGCCGCTGTATCGGAGAGGGGAGACCTTCCAGCCCACTGCAG 361
QY 602 GGTCTGTGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCGCTGTGCGAGTGAGGT 661
Db 362 CATCCGCTGCCGCTGCGAGGACGGCGCTTACCTGCCGTGCCGCTGTGCGAGGAGGT 421
QY 662 GCGGCTGCCAGCTGGGACTGCCCCACGCCCCACAGAGATACAGGTGCCAGGAAAGTGCTG 721
Db 422 GCGGCTGCCAGCTGGGACTGCCCCACGCCCCAGGAGGGTGCAGGTCTTGGGCAAGTGCTG 481
QY 722 CCCCAGTGGGTATGTGACAGGGAGTGACACGGCGCATCCAGCGCTCCACGGCGCAAGG 781
Db 482 CCCTGAGTGGGTGTGCGGCCAAGAGGGGGAGTGGGGACCCAGCCCTTCCAGCCCAAGG 541
QY 782 ACACCAACTTTCTGCCCTTGTCACTCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 841
Db 542 ACCCAGTTTCTGGGCTGTGTCTTCCCTGCCCTGCTGTGTGTGTGTGTGTGTGTGTGT 601
QY 842 CACAGCCTGGGCGCCCTGCTCAACCACTGTGGGCTGGGCAT 883
Db 602 CACGGCCTGGGACCTGCTCGACCACTGTGGGCTGGGCAT 643

RESULT 2
BQ937887 940 bp mRNA linear EST 21-AUG-2002
LOCUS
DEFINITION AGENCOURT_8951807 NCI_CGAP_Co24 Mus musculus cDNA clone
IMAGE:6476852 5', mRNA sequence.
ACCESSION BQ937887
VERSION BQ937887.1 GI:22353365

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14017 row: n column: 21
High quality sequence stop: 543.

FEATURES
Location/Qualifiers
1..940
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6476852"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 169 a 277 c 288 g 200 t 6 others
ORIGIN
Query Match 48.2%; Score 425.8; DB 14; Length 940;
Best Local Similarity 90.9%; Pred. No. 2.6e-98;
Matches 488; Conservative 0; Mismatches 39; Indels 10; Gaps 3;

QY 1 GACGCTTCTGATCCAGAGGACCCCTGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCC 60
Db 37 GACGCTCCTGATCCAGAGGACCCCGGCTGGGACAGGGGCTTGGCGAGGCTGCAGCT 96
QY 61 GCTG-GGCAGTGGCTTGAATGGAGGTCTTTTATTACTGGAACTGAGGAGCTAAGAGGCT 119
Db 97 GCTGTGGCAGTAGCTGGGATGGAGGTCTTTCTTCTGCTGGAACTGAGGAGCTGAGAAGCT 156
QY 120 CCTGTCTAG--CTTCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGT 176
Db 157 CCTGTAGGCTCCTCTCTAAACTCTTGGCACTTGGGCTTGGGCTTCACACACTGT 216
QY 177 CAGACACCTTCTGCTGGCTTCCACGGCTCCTACCTTCAAGTTTGAAGCTGGTCCACAAG 236
Db 217 CAGACACCTTCTGCTGGCTCCTCTCGGCC-----TCAGGTTTGAAGCTGGTCCACAAG 270
QY 237 GGACACGGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTCTGC 296
Db 271 GGACACGGTGACATGAGGGGCGAACCCTGATCCATCTTCTGGCCATTTCTCTCTCTGC 330
QY 297 CTCTCTCAATGGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTCTTGGACACCA 356
Db 331 ATTCTCTCAATGGTGTATGCCAGCTGTGCCAGACCCCTGTGCTCTCTCTTGGACACCA 390
QY 357 CCCCAGTGGCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTTAAAGTGTGT 416
Db 391 CCCCAGTGGCCACCGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTNAGTGTGT 450
QY 417 GCACGAGGCTGGGGAGTCTCTGGACCACTGCTGCGACCCAGCCAGGCGCTG 476
Db 451 GCACGAGGCTGGGGAGTCTCTGGACCACTGCTGCGACCCAGCCAGGCGCTG 510
QY 477 GTTTGTAGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAG 533

Db 511 GTTTGTCAGCCTGGGGCAGGCCCCCAGTGGCCGTTGGTGCTGTGTGCCCTCTGTGCTANG 567

RESULT 3

BQ279131 979 bp mRNA linear EST 07-MAY-2002

LOCUS AGENCOURT_7046721 NIH_MGC_107 Homo sapiens cDNA clone IMAGE:5805819

DEFINITION 5', mRNA sequence.

ACCESSION BQ279131

VERSION BQ279131.1 GI:20489339

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 979)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLCM2045 row: k column: 04

High quality sequence stop: 752.

Location/Qualifiers

1. .979

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5805819"

/clone_lib="NIH_MGC_107"

/tissue_type="adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: breast; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 141 a 347 c 313 g 170 t 8 others

ORIGIN

Query Match 47.5%; Score 419.8; DB 14; Length 979;

Best Local Similarity 80.2%; Pred. No. 9.1e-97;

Matches 493; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 269 CCATCTTCTGGCCACTTCCTTCCTCTGCGCTTCTCAATGGTGTGCGCCAGCTGTGCCG 328

Db 13 CCACCTCCTGGCCTTCTCCCTCCTCTGCGCTCCTCTCAAAGGTGCGTACCCAGCTGTGCC 72

QY 329 GACACCCCTGTACCTGTCTTGGACACACCCAGTGCACAGGGGTACCCCTGGTGCT 388

Db 73 GACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCGTGGAGTACCCCTGGTGCT 132

QY 389 GGATGGCTGTGGCTGTGTAAAGTGTGTCACGGAGGCTGGGGGAGTCTCTGCGACCACT 448

Db 133 GGATGGCTGTGGCTGTGCCGGGTATGTGCACGGCGGCTGGGGGAGCCCTGCGACCACT 192

QY 449 GCATGTCTGGACCCAGCCAGGCCCTGGTTGTACAGCCTGGGGGAGCCCTGGCGGCCA 508

Db 193 CCACGTCTGCGACGCCAGCCAGGCGCCTGGTCTGCCAGCCCGGGGAGGACCCGGTGACG 252

QY 509 TGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGTGAATGGCCGAGGTA 568

Db 253 GGGGGCCCTGTGCCTCTTTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGCCCGCCTGTA 312

QY 569 CCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTGTGCGCGTGTGATGACGGTGG 628

Db 313 TCGGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGG 372

QY 629 CTTACACCTGCTGCCGCTGTGTGCAGTGAGGATGTGCGGCTGCCAGTGGGACTGCCACG 688

Db 373 CTTACACCTGCTGCCGCTGTGTGCAGCGAGGATGTGCGGCTGCCAGTGGGACTGCCCCCA 432

QY 689 CCCCAAGAGATAACAGTGCAGGAAGTGTGCCCGAGTGGGTATGTGACCAAGGAGT 748

Db 433 CCCAGGAGGTCGAGGTCTCTGGGCAAGTGTGCCCTGAGTGGGTGTGCGGCGCAAGGAGG 492

QY 749 GACACCGCGGATCCAGCGCTCCACGGCGCAAGGACACCAACTTTCTGCCCTTGTCACTCC 808

Db 493 GGGACTGGGACCCAGCCCCCTTCCAGGCCCAAGGACCCAGTTTCTGGCCTTGTCTCTTC 552

QY 809 TGCCTCTGTGATGCTCCTTGTCCAAATTGGAGCACAGCCTGGGGCCCTGCTCAACCAC 868

Db 553 CCTGCCCCCTGGTGTCCCCTGCCAGATGGAGCACGGCCTGGGGACCCCTGCTCGACCAC 612

QY 869 CTGTGGGCTGGGCAT 883

Db 613 CTGTGGGCTGGGCAT 627

RESULT 4

BQ073722

LOCUS AGENCOURT_7046577 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:5806602

DEFINITION 5', mRNA sequence.

ACCESSION BQ073722

VERSION BQ073722.1 GI:19902768

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1073)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLCM2047 row: k column: 19

High quality sequence stop: 517.

Location/Qualifiers

1. .1073

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5806602"

/clone_lib="NIH_MGC_101"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 166 a 384 c 347 g 176 t

ORIGIN

Query Match 46.8%; Score 413; DB 14; Length 1073;

Best Local Similarity 80.2%; Pred. No. 5.2e-95;

Matches	485;	Conservative	0;	Mismatches	120;	Indels	0;	Gaps	0;
QY	276	CTGGCCACTTCCCTTCCCTCTGCTCCCTTCTCTCAATGGTGTGTGCCCCAGCTGTGCGGACACCC	335						
Db	2	CTGGCCCTTCTCCCTCTCTGCTCCCTCTCTCAAGGTGCGTACCCAGCTGTGCCCCGACACCA	61						
QY	336	TGTACCTGTCTCTTGGACACACACCCAGTGCCTCCACAGSGGTACCCCTGGTGTGCTGGATGSC	395						
Db	62	TGTACCTGCCCCGTGGCCACCTCCCCGATGCCCGCTGGAGTACCCCTGGTGTGCTGGATGSC	121						
QY	396	TGTGGCTGCTGTAAAGTGTGTGCACGGAGGCTGGGGGAGTCTCTGCGACCACTGCATGTC	455						
Db	122	TGTGGCTGCTGCCGGGTATGTGCACGGCGGCTGGGGGAGCCCTGCGACCACTCCACGTC	181						
QY	456	TGCGACCCACGCCAGGGCCCTGGTTGTGACGCCCTGGGCGAGGCCCTGGCGGGCCATGGGGCT	515						
Db	182	TGCGACGCCACGCCAGGGCCCTGGTCTGCCAGCCCGGGCGAGGACCCGGTGGCCGGGGGCC	241						
QY	516	GTGTGTCTCTTGGATGAGGATGACGGTAGCTGTAGTGAATGGCCGACAGTACCTGGAT	575						
Db	242	CTGTGCCCTCTTGGCAGAGGACGACAGCAGCTGTAGTGAACGGCCGCTGTATCGGGAA	301						
QY	576	GGAGAGACCTTTAAACCCCAATTGCAGGGTCCCTGTGCCGCTGTGATGACGGTGGCTTCACC	635						
Db	302	GGGAGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGGCTTCACC	361						
QY	636	TGCCTGCCGCTGTGCAGTGAGGATGTGCGGCTGCCAGCTGGGACTGCCACAGCCCCAAG	695						
Db	362	TGCGTGCCGCTGTGCAGCGAGGATGTGCGGCTGCCAGCTGGGACTGCCCCCAACCCCAAG	421						
QY	696	AGAATACAGGTGCACGGAAGTGTGCCCCGAGTGGGTATGTGACCAGGGAGTGACACCG	755						
Db	422	AGGTCGAGGTCTTGGGCAAGTGTGCCCTGAGTGGGTGTGCGGCCAAGGAGGGGGACTG	481						
QY	756	GCGATCCAGCGCTCCACGGCGCAAGGACACCACTTCTTGCCCTTGTCTACTCCTGCCCT	815						
Db	482	GGGACCCAGCCCTTCCAGCCCCAAGGACCCCAAGTTTCTTGCCCTTGTCTCTTCCCTGCC	541						
QY	816	GCTGATGCTCTTGTCTCCAAATTGGAGCACAGCTGGGGCCCCCTGCTCAACACCTGTGGG	875						
Db	542	CCTGGTGTCCCTTGCCACAGAATGGAGCACGGGCTGGGACCCCTGCTCGACCACCTGTGGG	601						
QY	876	CTGGG	880						
Db	602	CTGGG	606						

RESULT	5
BM043988	
LOCUS	750 bp mRNA
DEFINITION	BM043988 linear EST 07-NOV-2001 603620978F1 NTH_MGC_40 Homo sapiens cDNA clone IMAGE:5446794 5', mRNA sequence.

ACCESSION BM043988
VERSION BM043988.1 GI:16773255
KEYWORDS EST.
SOURCE human.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 750)
AUTHORS	NIH-MGC http://mgc.ncbi.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, ph.D.

RESULT 6
BM921531
LOCUS
DEFINITI

BM921531	1006 bp	mRNA	linear	EST 12-MAR-2002
AGENCOURT 6708025	NIH MGC 115	Homo sapiens	CDNA clone	IMAGE:5753009

FEATURES	source
High quality sequence stop: 714.	
Location/Qualifiers	
1. .750	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:5446794"	
/clone_lib="NIH_MGC_40"	
/tissue_type="carcinoma, cell line"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: prostate; Vector: pOTB7; Site_1: xhoI;	
Site_2: EcoRI; cDNA made by oligo-dT priming.	
Directionally cloned into EcoRI/xhoI sites using the	
following 5' adaptor: GGCACGAG(G). Library constructed by	
Ling Hong in the laboratory of Gerald M. Rubin (University	
of California, Berkeley) using ZAP-cDNA synthesis kit	
(Stratagene) and Superscript II RT (Life Technologies).	
Note: this is a NIH_MGC Library."	
BASE COUNT	112 a 267 c 246 g 125 t
ORIGIN	

BASE COUNT	112 a	267 c	246 g	125 t
ORIGIN				

Query Match 45.4%; Score 400.6; DB 13; Length 750;
Best Local Similarity 79.5%; Pred. No. 6.6e-92;
Matches 511; Conservative 0; Mismatches 129; Indels 3; Gaps 3;

QY	243	GGTGACATGAGGGCAGCCCACTGATCCATCTTCTGGCCCACTTCCTTCCCTTGCCTTCTC	302
DB	10	GGTGACATGAGAGGCACACCGAAGACCCACCTCTCTGGCCTTCTCCCTCCTCTGCCTCCTC	69
QY	303	TCAATGGTGTGTCCACAGTGTGCCGGACACCGTGTACCTGTCTCTGGACACCAACCCACG	362
DB	70	TCAAAGGTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCCACTCCCCGA	129
QY	363	TGCCCCACAGGGGTACCCCTGGTGGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGG	422
DB	130	TGCCCGCTGGGAGTACCCCTGGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG	189
QY	423	AGGCTGGGGAGGCTCGCAGCACCACTGCATGTCTGCGACCCCA - GCCAGGGCCTGGTTTG	481
DB	190	CGGCTGGGGAGGCTCGCAGCAACTCCACGTCTGCGACGCCATGCCAGGGCCTGGTCTG	249
QY	482	TCAGCCTGGGCGAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGG	541
DB	250	CCAGCCCGGGCAGGACCCGGTGGAGCGGGGGCCCTGTGCCCTCTTGGCAGAGGACGACAG	309
QY	542	TAGCTGTGAGGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG	601
DB	310	CAGCTGTGAGGTGAACGGCCCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCACTGCAG	369
QY	602	GGTCCTGTGCCGGTGTGATGACGGTGGCTTCACCTGCCCTGCCCTGTGCACTGAGGATGT	661
DB	370	CATCCGCTGCCGGTGGGAGGACGGCGGCTTCACCTGCGTGCCGCTGTGCAAGGATGT	429
QY	662	GCGGCTGCCAGTGGGACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTGTCTG	721
DB	430	GCGGCTGCCAGTGGGACTGCCCCCAACCCAGGAGGGTTCGAGGTCTCTGGCAAGTGTCTG	489
QY	722	CCCCGAGTGGGTATGTGACCAAGGAGTGACACCGCGGATCC - AGCGCTCCACGGCGCAAG	780
DB	490	CCCTGAGTGGGTGTGCGGCCAAGGAGGGGGACTGGGGACCCCAAGCCCCCTTCAGGCCAAG	549
QY	781	GACACCAACTTTGTGCCCCCTGTCACTCCTGCTGTGCTGATGCTCCTTGTCCAAATTTGA	840
DB	550	GACGCCAGTTTTCTTGGCCTTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCCAGAAATGGA	609
QY	841	GCACAGCCTGGGGCCCCCTGCTCAACCAACCTGTGGGCTGGGCAT	883
DB	610	GCACGGCCTGGGGACCCCTGCTCGA - CACCTGTGGGCTGGGCAT	651

[illegible]

QY 315 GCCAGCTGTGCCGGACACACCTGTACCTGTCTCTTGGACACACACCCACCTGCCCCACAGGGG 374
|||||
Db 61 ACCAGCTGTGCCCGACACACCTGTACCTGTCTCTTGGCCACCTCCCGATGCCCGCTGGGA 120

QY 375 GTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAAGTGTGTGACGAGGCTGGGGGAG 434
|||||
Db 121 GTACCCCTGGTGTGGATGGCTGTGGCTGTCTGCCGGGTATGTGACGGCGCTGGGGGAG 180

QY 435 TCCTCGGACCACTGCATGTCTCGACCCCGACCCAGCGGCGCTGTGTTGTGACGCTGGGGCA 494
|||||
Db 181 CCCTCGGACCACTCCACGCTCTCGACGCCAGCCAGGCGCTGTCTGCCAGCCCGGGGCGAG 240

QY 495 GGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTG 554
|||||
Db 241 GACCTGGTGGACGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCTGTGAGGTG 300

QY 555 AATGGCCGCGAGGTACCTGGATGGAGAGACCTTTAAACCAATTCAGGGTCTGTGCGGC 614
|||||
Db 301 AACGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCGACTGCAGCATCCGCTGCCGC 360

QY 615 TGTGATGACGGTGGCTTACCTGCTTCCCGCTGCGCTGTGTCAGTGAGGATGTGCGGCTGCCAGC 674
|||||
Db 361 TGCAGGACGGCGCTTACCTGCTGCGCTGTGTCAGGAGGATGTGCGGCTGCCAGC 420

QY 675 TGGACTGCCACGCCCGCCAGAGAAATACAGGTGCCAGGAAAGTGTGCGCCCGAGTGGSTA 734
|||||
Db 421 TGGACTGCCCGCCACCCAGAGGGTGCAGGTCTTGGGCAAGTGTGCTGCTGAGTGGGTG 480

QY 735 TGTGACGAGGAGTGACACCGCGGATCCAGCGCTCCAGCGCGCAAGGACACCACTTCT 794
|||||
Db 481 TGCG-GCAAGAGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGGACCCAGTTTCT 539

QY 795 GCCCTGTCACTCCTGCTGTGATGCTCCTTGTTCAAATGGAGCACAGCCTGGGGC 854
|||||
Db 540 GGCTTGTCTCTTCCCTGCCCGCTGGTGTCCCTGCCAGAAATGGAGCACGGCCTGGGGA 599

QY 855 CCCTGCTCAACCACTGTGGCTGGGCAT 883
|||||
Db 600 CCTGCTCGACCACCTGTGGCTGGGCAT 628

RESULT 8
BI457141
LOCUS
DEFINITION BI457141 651 bp mRNA linear EST 21-AUG-2001
603185392F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5258159 5',
mRNA sequence.
ACCESSION BI457141
VERSION BI457141.1 GI:15247797
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 651)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1869 row: g column: 24
High quality sequence stop: 651.
FEATURES
source Location/Qualifiers
1..651
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:5258159"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: This is a NIH_MGC Library. |"
BASE COUNT 91 a 230 c 219 g 111 t

Query Match 40.4%; Score 391.8; DB 13; Length 651;
Best Local Similarity 79.5%; Pred. No. 1.1e-89;
Matches 476; Conservative 0; Mismatches 122; Indels 1; Gaps 1;

QY 285 TCCTTCCTCTGCCCTCTCTCAATGGTGTGTGCCAGCTGTGCGCGGACACCTGTACCTGT 344
|||||
Db 6 TCCTTCCTCTGCCCTCTCTCAAGGTGCGTACCCAGCTGTGCGCGGACACCTGTACCTGC 65

QY 345 CTTGGACACACCGCCAGTCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGC 404
|||||
Db 66 CCTGGCCACCTCGCCGATGCCGCTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGC 125

QY 405 TCTAAAGTGTGTGACGGAGGCTGGGGAGTCTCTCGACACCTGTCATGTCTCGACCCC 464
|||||
Db 126 TGCCGGGTATGTGACGGCGCTGGGGAGCCCTCGGACCAACTCCACGTCTCGGACGCC 185

QY 465 AGCCAGGGCCCTGGTCTGTCAAGCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTC 524
|||||
Db 186 AGCCA-GGCTGGTCTGCCAGCCCGGGCAGGACCCCGTGGACGGGGCCCTGTGCTC 244

QY 525 TTGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCCGAGGTACCTGGATGGAGAGACC 584
|||||
Db 245 TTGGCAGAGGACGACAGCAGTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACC 304

QY 585 TTTAAACCCCAATTGCAAGGTCTCTGCCGCTGTGATGACGGTGGCTTCACCTGCCCTGCCG 644
|||||
Db 305 TTCCAGCCCCACTGCAAGCATCCGCTGCCGCTGCGAGGACGGCGGCTTCACCTGCCGTGCCG 364

QY 645 CTGTGCAGTGAGGATGTGGGCTGCCAGCTGGGAGTGGCCACGCCACGCCCAAGAGATACAG 704
|||||
Db 365 CTGTGCAGCGAGGATGTGGGCTGCCAGCTGGGACTGCCCCCAACCCAGGAGGCTCGAG 424

QY 705 GTGCCAGGAAAGTGTGCCCGAGTGGGTATGTACCAGGGAGTGACACCGCGGATCCAG 764
|||||
Db 425 GTCCTGGGCAAGTGTGCCCTGAGTGGGTGTGCGGCCAAGAGGGGACTGGGGACCCAG 484

QY 765 CGCTCCACGGCGCAAGGACACCAACTTTCTGCCCTTCTCACTCTCTGCTGCTGATGCT 824
|||||
Db 485 CCCCTTCCAGCCCAAGGACCCAGTTTCTGGCCCTTCTCCTCTTCCCTGCCCTGCCCTGTGT 544

QY 825 CCTTGTCCAAATTGGACACAGCCCTGGGGCCCTGCTCAACACCTGTGGGCTGGGCAT 883
|||||
Db 545 CCCTGCCCCAGATGGACGACGGCCTGGGGACCCCTGCTCGACACCTGTGGGCTGGGCAT 603

RESULT 9
BQ560868
LOCUS
DEFINITION BQ560868 537 bp mRNA linear EST 20-JUN-2002
H4067A01-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
H4067A01 5', mRNA sequence.
ACCESSION BQ560868
VERSION BQ560868.1 GI:21461753
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

Db 382 TCGGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGTGCGAGGACGGCGG 441

QY 629 CTTACACCTGCCCTGCTGCAGTGAGGATGTGCGGCTGCCAGCTGGGACTGCCACG 688

Db 442 CTTACACCTGCCCTGCTGCAGCGAGGATGTGCGGCTGCCAGCTGGGACTGACCCCA 501

QY 689 CCCCAGAGATACAGGTGCCAGGAAGTGTGCTGCCCGAGTGGGTATGTGACCAAGGAGT 748

Db 502 ACCAGGAGGTCGAGGTCTTGGCAAGTGTGCTGCCCTGAGTGGGTGTGCGGCCAAGGAGG 561

QY 749 GACACCGGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTTCGCCCTTGTCACT-- 806

Db 562 GGGACTGGGACCCAGACCCCTTACCAGCCCAAGGACCCACAGATTCTGGCCTTGTCTTTC 621

QY 807 CCTGCCCTCTGCTGATGCTCCTTGTCCAAATTTGGAGCACAGCCTGGG-CCCCTGCTCAAC 865

Db 622 CCATGCCCCACTGGTGTCCCTTGACCAAGATGGAGCACGGCTTGGGAAACCCCTGCTCGAC 681

QY 866 CACCTGTGGGCTGGG 880

Db 682 CACATGTGGGCTGG 696

RESULT 11

BI826781

LOCUS

DEFINITION

603077268F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169065 5',

mRNA sequence.

ACCESSION

BI826781

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 800)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM11420 row: g column: 18

High quality sequence stop: 788.

FEATURES

source

1..800

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5169065"

/clone_lib="NIH_MGC_119"

/tissue_type="medulla"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

BASE COUNT

115 a 284 c 264 g 136 t 1 others

ORIGIN

Query Match 41.6%; Score 367.2; DB 13; Length 800;

Best Local Similarity 79.0%; Pred. No. 2.4e-83;

Matches 486; Conservative 0; Mismatches 124; Indels 5; Gaps 4;

QY 269 CCATCTTCTGGCCACTTCTTCTCCTCTGCTTCTCAATGGTGTGTGCCAGCTGTGCCG 328

Db 14 CCACCTCCTGGCCTTCTCCTCCTCTGCTCCTCTCTCAAAGGTGCGTACCAGCTGTGCC 73

QY 329 GACACCCCTGTACCTGTCTTGGACACCAACCCAGTCCCCACAGGGGGTACCCCTGGT 388

Db 74 GACACCATGTACGTGCCCTGCCACCTCCCCGATGCCCGTGGGAGTACCCCTGGTGT 133

QY 389 GGATGGCTGTGGTGTCTGTAAAGTGTGTGCACGGAGGCTGGGGAGTCTCTGCGACCA 448

Db 134 GGATGGCTGTGGTGTCTGCGGGTATGTGCACGGCGGTGGGGAGCCCTGCGACCACT 193

QY 449 GCATGTCTGCGACCCAGCCAGCCAGGCGCTGGTTTGTACGCTGGGGAGGCCCTGGCG 508

Db 194 CCACGTCTGCGACGCCAGCCA-GGCCTGGTCTGCCAGCCCGGGCAGGACCCGGTGG 252

QY 509 TGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGTGAATGGCCGAGGTA 568

Db 253 GGGGGCCCTGTGCTCTTGGCAGAGGAGACAGCAGCTGTGAGTGAACGGCCGCTGTA 312

QY 569 CCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGGTCTCTGCCCTGTGATGACGGT 628

Db 313 TCGGGAAGGGAGACCTTCCAGCCCACTGCAGCATCCGCTGCCGTGCGAGGACGGCG 372

QY 629 CTTACCTCCTGCCCTGCTGCAGTGAGGATGTGCGGCTGCCAGCTGGGACTGCCACG 688

Db 373 CTTACCTCCTGCCCTGCTGCAGCGAGGATGTGCGGCTGCCAGCTGGGACTGCCCCA 432

QY 689 CCCCAGAGAAATACAGGTGCCAGGAAGTGTGCCCCGAGTGGGTATGTGACCAAGGAG 748

Db 433 CCCAGAGGGTCGAGGTCTTGGCAAGTGTGCCCTGAGTGGGTGTGCG-GCAAAGGAG 491

QY 749 GACACCGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCCCTTGTCACT 808

Db 492 GGGACTGGGGACCCAGCCCTTCCAGCCCAAGGACCCAGTTTCTGGCCTTGTCTCTTC 551

QY 809 TGCCTCTGCTGATGCTCCTTGTCCAAATTTGGAGCACAGCCTGGGCGCCCTGCTCA 868

Db 552 -CCTGCCCTGGTGTCCCTGCCAGATGGAGNCACGGC--TGGGACCCCTGCTCGACC 608

QY 869 CTGTGGGCTGGGCAT 883

Db 609 CTGTGGGCTGGGCAT 623

RESULT 12

BI825652

LOCUS

DEFINITION

603072631F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5164614 5',

mRNA sequence.

ACCESSION

BI825652

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 888)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM11408 row: n column: 07

High quality sequence stop: 877.

FEATURES source Location/Qualifiers

1. .888

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5164614"

/clone_lib="NIH_MGC_119"

/tissue_type="medulla"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

BASE COUNT 130 a 308 c 288 g 162 t

ORIGIN

Query Match 41.6%; Score 367.2; DB 13; Length 888;

Best Local Similarity 77.1%; Pred. No. 2.5e-83;

Matches 472; Conservative 0; Mismatches 138; Indels 2; Gaps 2;

QY 273 CTTCTGGCCACTTCCTCTCCTCTCCTCTCAATGGTGTGTCGCCAGCTGTGCCGGACA 332

Db 15 CCTCTGGCTTCCTCCTCTCCTCTCTCAAGGTGCGTACCCAGCTGTGCCCGACA 74

QY 333 CCCTGTACCTGTCTTGGACACACCCAGTGCCACAGGGGTACCCGTGGTGGAT 392

Db 75 CCATGTACCTGCCCTGGCCACCTCCCGATGCCGCTGGGAGTACCCCTGGTGGAT 134

QY 393 GGCTGTGGCTGTGTAAGTGTGTGCACGGAGGCTGGGGAGTCTCTGCACCACTGCAT 452

Db 135 GGCTGTGGCTGTGCCGGGTATGTGCACGGCGGCTGGGGAGCCCTGGCACCACTCCAC 194

QY 453 GTCTGCGACCCAGCCAGGGCCCTGGTTGTGACGCTGGGGCAGGCCCTGGCGCCATGGG 512

Db 195 GTCTGCGACGCCAGCCAGGCCTGTGTCTGCCAGCCCGGGCAGGACCCGGTGGCGGGG 254

QY 513 GCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAATGCCCGCAGGTACCTG 572

Db 255 GCCCTGTGCCTTGGCAGAGGACGACAGAGCTGTGAGGTGAACGGCCGCTGTATCGG 314

QY 573 GATGGAGAGACCTTTAAACCCCAATTGCAGGGTCCTGTGCCGCTGTGATGACGGTGGCTTC 632

Db 315 GAAGGGAGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGCTTC 374

QY 633 ACCTGCCTGCCGCTGTGCAGTGAGGATGTGGGCTGCCAGCTGGGACTGCCACGCCCC 692

Db 375 ACCTGCGTGCCGCTGTGCAGCGAGGATGTGGGCTGCCAGCTGGGACTGCCCCACCCC 434

QY 693 AAGAGAATACAGTGCCAGGAAAGTGTGCCCGGAGTGGGTATGTGACGAGGAGTGACA 752

Db 435 AGGAGGTCGAGGTCTTGGCAAGTGTCTGCCCTGAGTGGGTGTGGGCAAGGAGGGGA 494

QY 753 CCGGCGATCCAGCGCTCCACGGCGCAAGSACACCAACTTTCTGCCCTTGTTCAC-TCCTGC 811

Db 495 CTGGGACCCAGCCCTTCCAGCCCCAAGSACCCAGTTTCTGGCCTTGTCTCTTCCCTG 554

QY 812 CTCTGTGATGCTCTTGTCCAAATTGGACACAGCCCTGGGGCCCTGTCAACCACTG 871

Db 555 GCCCCCTGGTGTCCCTGTCCCAAGATGGAGCAC-GGCTGTGGACCCCTGCTCGACCACCTG 613

QY 872 TGGGCTGGGCAT 883

Db 614 TGGGCTGGGCAT 625

RESULT 13

BG928868

LOCUS

DEFINITION HNC57-1-D9.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA

sequence.

ACCESSION BG928868

VERSION BG928868.1 GI:14323391

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 620)

AUTHORS Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J., Sathe,G., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and Lark,M.W.

TITLE Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries

JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)

MEDLINE 21482651

COMMENT Contact: Sanjay Kumar UW2109 GlaxoSmithKline 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA Tel: 610-270-7245 Fax: 610-270-5598 Email: sanjay_kumar-l@gsk.com Seq primer: T7.

FEATURES source Location/Qualifiers

1. .620

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HNC (Human Normal Cartilage)"

/tissue_type="cartilage"

/lab_host="E.coli DH10 B"

/note="Vector: pSPORT I; Site_1: SalI; Site_2: NotI; Directional"

BASE COUNT 97 a 218 c 207 g 98 t

ORIGIN

Query Match 41.4%; Score 366; DB 13; Length 620;

Best Local Similarity 79.7%; Pred. No. 4.4e-83;

Matches 432; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 244 GTGACATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCTCTCCTCTCTCT 303

Db 79 GCGTCCGGAGAGGCACACCGAAGACCCACCTCTCGCCTTCTCCTCTCTCTCTCTCT 138

QY 304 CAATGGTGTGTCGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCACTCC 363

Db 139 CAAAGTGTGCTACCCAGCTGTGCCCGACACCATGTACCTGCCCTTGGCCACCTCC 198

QY 364 GCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAGTGTGTGACCGGA 423

Db 199 GCCCGCTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGCGGGTATGTGACGGC 258

QY 424 GGCTGGGGAGTCTCGACCACTGCATGTCTGCGACCCAGCCAGGGCCCTGGTTTGTG 483

Db 259 GGCTGGGGAGCCCTCGACCAACTCCACGTCTGCGACGCCAGCGAGGGCCTGGTCTGCC 318

QY 484 AGCCTGGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGTA 543

Db 319 AGCCCGGGCAGGACCCGGTGGCCGGGGGCCCTGTGGCTCTTGGCAGAGGACGACAGCA 378

QY 544 GCTGTGAGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGG 603

Db 379 GCTGTGAGTGAACGGCCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGCA 438

QY 604 TCCTGTGCCGCTGTGATGACGGTGGCTTCACTGCCTGCCCTGTGTCAGTGAGGATGTGC 663

Db 439 TCCGCTGCCGCTGCGAGGACGGCGCTTCACTGCCTGCCGCTGTGTCAGCGAGGATGTGC 498

QY 664 GGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGAATACAGGTGCCAGGAAAGTGTGCC 723

Db 499 GGCTGCCAGCTGGGACTGCCCCCAAGGAGGTCGAGGTCTCGAGGTCTCGGCAAGTGTGCC 558

constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT 216 a 424 c 315.g 207 t 4 others
ORIGIN

Query Match 40.9%; Score 361.2; DB 13; Length 1166;
Best Local Similarity 78.8%; Pred. No. 9.9e-82;
Matches 458; Conservative 0; Mismatches 114; Indels 9; Gaps 2;

Qy	167	CACACACTGT	CAGACACCTTC	CGTGGTGGCCCT	CCACGGCCCTCACCTTC	CAGGTTTGAAGCTG	226
Db	137	CACACACGGAC	AGGCACCCCTT	GGTGGCCTT	CACAGTTTCACCTTC	CAGGCTCAAAGCTG	196
Qy	227	GCTCCACAAGG	GACACGGT	GACATGAGGGG	CAGCCCACTGATCCATCTT	CTGGCCACTTC	286
Db	197	GCTCTGCAGGG	-----G	ACATGAGAGGG	CACACCCGAGACCCACCTC	CTGGCCTTCTC	248
Qy	287	CTTCCTCTGCC	TTCTCTCAAT	GGTGTGTGCC	CAGCTGTGCCGGACACCCCTG	TACCTGTCC	346
Db	249	CTCCTCTGCC	TTCTCTCAAG	GTGCGTACCC	CAGCTGTGCCGGACACCATG	TACCTGCCC	308
Qy	347	TTGGACACCA	CCCCAGTGC	CCACAGGGGGT	ACCCCTGGTGGATGGCTG	TGGCTGTG	406
Db	309	CTGCCACCTC	CCCCGATG	CCCGCTGGGAGT	ACCCCTGGTGGATGGCTG	TGGCTGTG	368
Qy	407	TAAAGTGTGT	GCACGGAGG	CTGGGGGAGT	CTCTGCGACCACTGCATGT	CTGCGACCCAG	466
Db	369	CCGGTATGTG	CACGGCGG	CTGGGGAGCC	CTGGGACCAACTCCACGT	CTGCGACGCCAG	428
Qy	467	CCAGGGCCTG	GTTCAGCT	TGGGGCAGGCC	CTGGCGGCATGGGGCTGT	TGTCTCTT	526
Db	429	CCAGGGCCTG	GTCTGCCAG	CCCCGGGCGAGG	ACCCGGTGGACGGGGGCCCT	TGTGCCTCTT	488
Qy	527	GGATGAGGAT	GACGGTAG	CTGTGAGGTGA	ATGGCCGAGGTACCTGGAT	GGAGAGACCTT	586
Db	489	GCGAGGACG	ACAGCAGCT	GTGAGGTGA	ACGGCCCTGTATCGGGA	AGGGGAGACCTT	548
Qy	587	TAAACCCAA	TTCAGGGT	CTCTGTGCCG	CTGTGATGACGGTGGCTT	CACCTGCCCTG	646
Db	549	CCAGCCCCA	CTGCAGCA	TCCGCTGCCG	CTGCGAGGACGCGGCTT	CACCTGCCGTGCC	608
Qy	647	GTGCAGTGA	GAGTGTG	CGGCTGCC	CAGCTGGGACTGCCAC	-GCCCCAAGAGA	705
Db	609	GTGCAGCGA	AGATGTG	CGGCTGCC	CAGCTGGGACTGCC	CCCCACCC	668
Qy	706	TGCCAGGAA	AGTGCTG	CCCCCAGT	GGGTATGTG	ACCAGGGA	746
Db	669	TCCTGNGCA	AGTGCTG	CCCCCTGA	ATGGGTGTG	CGGCCCAAGA	709

Search completed: July 29, 2003, 02:51:26
Job time : 1569.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2003, 01:19:19 ; Search time 2638 Seconds
(without alignments)
2758.037 Million cell updates/sec

Title: US-10-010-408-2
Perfect score: 1440
Sequence: 1 MRGSPLIHLHLATSFLLLSM.....LCLPRPCLAAARSHSSWNSAF 250

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues 4109280

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10010408/runat_25072003_101855_17924/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cd1 -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10010408@cgn_l1_2496@runat_25072003_101855_17924 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description	
1	1390	96.5	1741	10	AF259981	AF259981 Rattus no	
2	1308.5	90.9	1734	6	AR210324	AR210324 Sequence	
c	3	1308.5	1734	6	AR210325	AR210325 Sequence	
	4	1308.5	90.9	1734	10	AF100778	AF100778 Mus muscu
	5	1308.5	90.9	1739	10	AF126063	AF126063 Mus muscu
6	1064	73.9	1266	6	AX076919	AX076919 Sequence	
7	1064	73.9	1266	6	AX464186	AX464186 Sequence	
8	1064	73.9	1283	9	AF083500	AF083500 Homo sapi	
9	1064	73.9	1293	6	AR210322	AR210322 Sequence	
c	10	1064	1293	6	AR210323	AR210323 Sequence	
	11	1064	1309	9	AF074604	AF074604 Homo sapi	
	12	1064	1427	9	AF100780	AF100780 Homo sapi	
13	1064	73.9	1450	9	BC017782	BC017782 Homo sapi	
14	1047	72.7	738	6	AR210337	AR210337 Sequence	
15	1045.5	72.6	841	6	AR210338	AR210338 Sequence	
16	575.5	40.0	1585	5	NVI271167	AJ271167 Notophtha	
17	564	39.2	2075	6	AR018957	AR018957 Sequence	
18	564	39.2	2075	6	AR118595	AR118595 Sequence	
19	564	39.2	2075	6	AR130891	AR130891 Sequence	
20	564	39.2	2075	6	AX206708	AX206708 Sequence	
21	564	39.2	2075	6	I11636	I11636 Sequence 1	
22	564	39.2	2075	6	I32210	I32210 Sequence 1	
23	564	39.2	2075	9	HUMCONGRO	M92934 Homo sapien	
24	564	39.2	2312	9	HSCTGF	X78947 H.sapiens m	
25	564	39.2	2998	6	AR119211	AR119211 Sequence	
26	564	39.2	2998	6	AR151276	AR151276 Sequence	
27	553	38.4	2190	5	AF463517	AF463517 Gallus ga	
28	553	38.4	2288	5	GGA298335	AJ298335 Gallus ga	
29	550.5	38.2	1953	5	XL043524	U43524 xenopus lae	
30	550.5	38.2	2267	6	AX206706	AX206706 Sequence	
31	550.5	38.2	2267	10	MUSFISP12B	M70642 Mouse FISP-	
32	550.5	38.2	2330	10	MUSTGFB	M80263 Mouse mRNA	
33	550.5	38.2	2334	10	BC006783	BC006783 Mus muscu	
34	550.5	38.2	2345	6	AX402019	AX402019 Sequence	
35	550.5	38.2	2345	10	AF120275	AF120275 Rattus no	
36	548.5	38.1	2338	6	E37595	E37595 Monoclonal	
37	548.5	38.1	2338	10	AB023068	AB023068 Rattus no	
38	542.5	37.7	2350	6	AR194011	AR194011 Sequence	
39	542.5	37.7	2350	6	AR201286	AR201286 Sequence	
40	541.5	37.6	1975	6	A28405	A28405 Sequence I	
41	541.5	37.6	1975	6	A28444	A28444 nov mRNA se	
42	541.5	37.6	1976	5	GGNOVMRNA	X59284 G.gallus no	
43	541	37.6	1805	5	CHKCEF	J04496 Chicken CEF	
c	44	537	137964	2	AC126895	AC126895 Rattus no	
	45	533	1496	4	SSU83916	U83916 Sus scrofa	

ALIGNMENTS

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
Db 377 CCGGGGTACCCCTGGTGGATGGCTGGTGGCTGTCGAGCTGTGCACGGAGGCTG 436
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 437 GGGAGTCTCGCACCACTGCATGCTGCGACCCAGCCAGGCGCTGGTTTGTGAGCCT 496
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 497 GGGGCAGCCCCAGTGGCCGTGGTGGTGGCTCTTCGAAGAGGATGACGGGAGCTGT 556
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db 557 GAGGTGAATGGCCGCAGGTACCTGGATGGGAGACCTTTAAACCAATTCAGGGTTTG 616
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 617 TGCCGCTGTGATGACGGTGGTTTCACTGCCTGCCGTGCCGTGTGAGGATGTGCGGCTG 676
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 677 CCCAGCTGGACTGCCACGCCCGCCAGGAGAATACAGGTGCCAGGAAGGTGCTGCCCGAG 736
QY 161 TrpValCysAspGlnGlyVal---ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179
Db 737 TGGGTGTGTGACCAGGCAGTGATGCAGCGCGCAATCCAGCCCTCTCAGCCCAAGGACAC 796
QY 180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199
Db 797 CAACCTTCTGCCCTTGTCACTCCTGCATCTGCCGATGGCCCTGTCCAAACTGGAGCACA 856
QY 200 AlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln 219
Db 857 GCCTGGGGCCCTGCTCAACCACTGTGGTGGGATAGCCACCGAGTATCCAACACAG 916
QY 220 AsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
Db 917 AACCGATTCTGCCAACTGGAGATCCAGCTGCGCTGTGTCTGTCCAGACCTGCCTGGCA 976
QY 240 AlaArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 977 TCCAGGAGCCACGGCTCATGGAACAGTGCCTTC 1009

RESULT 3
AR210325/c
LOCUS AR210325 1734 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 18 from patent US 6387657.
ACCESSION AR210325
VERSION AR210325.1 GI:21512526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 18 14-MAY-2002;
FEATURES Location/Qualifiers
source 1..1734
BASE COUNT 393 a 495 c 491 g 355 t
ORIGIN /organism="unknown"

Alignment Scores:
Pred. No.: 3.97e-67 Length: 1734
Score: 1308.50 Matches: 226
Percent Similarity: 93.63% Conservative: 9
Best Local Similarity: 90.04% Mismatches: 15
Query Match: 90.87% Indels: 1
DB: 6 Gaps: 1

US-10-010-408-2 (1-250) x AR210325 (1-1734)
QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db 1478 ATGAGGGGCAACCCACTGATCCATCTTCTGGCCATTTCCCTCTCTGCATTTCTCAATG 1419
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
Db 1418 GTGTATTTCCAGCTGTGCCAGCACCTGTGCCTGTCTTGGACACACCCAGTGCCCA 1359
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
Db 1358 CCGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTGTCAGTGTGCACGGAGGCTG 1299
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 1298 GGGGAGTCTGCGACCACTGCATGCTGCGACCCAGCCAGGCGCTGGTTTGTGAGCCT 1239
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 1238 GGGCAGGCCCCAGTGGCCGTGGTGTGTGTGCTCTTCGAAGAGGATGACGGAGCTGT 1179
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db 1178 GAGGTGAATGGCCGCAGGTACCTGGATGGGAGACCTTTAAACCAATTCAGGGTTTG 1119
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 1118 TGCCGCTGTGATGACGGTGGTTTCACTGCCTGCCGTGCCGTGTGAGGATGTGCGGCTG 1059
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 1058 CCCAGCTGGGACTGCCACGCCCGCCAGGAGAATACAGGTGCCAGGAAGGTGCTGCCCGAG 999
QY 161 TrpValCysAspGlnGlyVal---ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179
Db 998 TGGGTGTGTGACCAGGCAGTGATGCAGCGCGCAATCCAGCCCTCTCAGCCCAAGGACAC 939
QY 180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199
Db 938 CAACCTTCTGCCCTTGTCACTCCTGCATCTGCCGATGGCCCTGTCCAAACTGGAGCACA 879
QY 200 AlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln 219
Db 878 GCCTGGGGCCCTGCTCAACCACTGTGGTGGGATAGCCACCGAGTATCCAACACAG 819
QY 220 AsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
Db 818 AACCGATTCTGCCAACTGGAGATCCAGGCTGCGCTGTGTCTGTCCAGACCTGCCTGGCA 759
QY 240 AlaArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 758 TCCAGGAGCCACGGCTCATGGAACAGTGCCTTC 726

RESULT 4
AF100778 1734 bp mRNA linear ROD 17-DEC-1998
LOCUS AF100778
DEFINITION Mus musculus connective tissue growth factor related protein WISP-2 (Wisp2) mRNA, complete cds.
ACCESSION AF100778
VERSION AF100778.1 GI:4028578
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1734)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A., Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C., Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D., Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.
TITLE WISP genes are members of the connective tissue growth factor

family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors
Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)
99061933
PUBMED 9843955
2 (bases 1 to 1734)
Pennica,D.
AUTHORS Direct Submission
TITLE Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA
JOURNAL Location/Qualifiers
FEATURES 1. .1734
Source /organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="C57MG"
/cell_type="epithelial"
/tissue_type="mammary"
/note="transformed by Wnt-1"
1. .1734
/gene="Wisp2"
257. .1012
/gene="Wisp2"
/codon_start=1
/product="connective tissue growth factor related protein WISP-2"
/protein_id="AAC96320.1"
/db_xref="GI:4028579"
/translation="MRGNPLHLHLLAISFLCILSMVYSQICPAPCACPTWPPQCPGVP
LVLDGCGCCRVCCARRLGESCDHLHVCDPSQGLVCQPGAGPSGRGAVCLFEEDDGCEV
NGRRLDGETFKPNCRVLCRCDDGGFTCLPLCSEDVRLPSWDCPRPRRIQVPGRCPE
WVCDQAVMQPAIQPSSAQGHLSALVTPASADGPCPNWSTAWGPCSTTCGLGIATRV
NQNRFCQLEIQRRLLSRPCLASRSHGWSNAF"
BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN

Alignment Scores:
Pred. No.: 3.97e-67 Length: 1734
Score: 1308.50 Matches: 226
Percent Similarity: 93.63% Conservative: 9
Best Local Similarity: 90.04% Mismatches: 15
Query Match: 90.87% Indels: 1
DB: 10 Gaps: 1

US-10-010-408-2 (1-250) x AF100778 (1-1734)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
|||||:::|||||
Db 257 ATGAGGGGCAACCCACTGATCCATCTCTGGCCATTTCTCTCTCTCTCAATG 316
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrThrProGlnCysPro 40
|||:::|||||
Db 317 GTGTATTCACAGCTGTGCCAGCACCCCTGTGCCTGTCTTGGACACACCCAGTGCCCA 376
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeu 60
|||||
Db 377 CCGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTCTGTCTGTCTGTCTGTCTG 436
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
|||||
Db 437 GGGGAGTCTCTCGACCACTGTCATGTCTGCGACCCAGCCAGCGGCTGTGTGTCTGAGCCT 496
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
|||||
Db 497 GGGCAGGCCCCAGTGGCCGT 556
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
|||||
Db 557 GAGGTGATGGCCGACAGTACCTGGATGGGGAGACCTTTAAACCCCAATTGCAGGGTTTG 616
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
|||||
Db 617 TGCCGCTGTGATGACGGTGGTTTACCTGCCTGCCGCTGTGCAGTGAGGATGTGCGGCTG 676

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
|||||
Db 677 CCCAGCTGGACTGCCACGCCCCAGGAGAAATACAGGTGCCAGGAAGGTGCTGCCCGAG 736
QY 161 TrpValCysAspGlnGlyVal---ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179
|||||
Db 737 TGGGTGTGTGACAGGCAGTGATGCAGCGCGCAATCCAGCCCTCCTCAGCCCAAGGACAC 796
QY 180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199
|||||
Db 797 CAACCTTCTGCCCTTGTCACTCCTGCATCTGCCGATGGCCCTGTCCAAACTGGAGCACA 856
QY 200 AlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln 219
|||||
Db 857 GCCTGGGGCCCTGCTCAACCACCTGTGGTTGGGCATAGCCACCGAGTATCCAACCAG 916
QY 220 AsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
|||||
Db 917 AACCGATTCTGCCAACTGGAGATCCAGCGTGCCTGTGTCTGTCCAGACCCCTGCCTGGCA 976
QY 240 AlaArgSerHisSerSerTrpAsnSerAlaPhe 250
:::|||||
Db 977 TCCAGGAGCCACGGCTCATGGAAACAGTGCCTTC 1009

RESULT 5
AF126063
LOCUS AF126063 1739 bp mRNA linear ROD 12-OCT-1999
DEFINITION Mus musculus connective tissue growth factor-like protein precursor (Ctgfl) mRNA, complete cds.

ACCESSION AF126063
VERSION AF126063.1 GI:4337059
KEYWORDS

SOURCE Mus musculus.
ORGANISM Mus musculus,
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1739)
AUTHORS Kumar,S., Hand,A.T., Connor,J.R., Dodds,R.A., Ryan,P.J., Trill,J.J., Fisher,S.M., Nuttall,M.E., Lipshutz,D.B., Zou,C., Hwang,S.M., Voeta,B.J., James,I.E., Rieman,D.J., Gowen,M. and Lee,J.C.
TITLE Identification and cloning of a connective tissue growth factor-like cDNA from human osteoblasts encoding a novel regulator of osteoblast functions

J. Biol. Chem. 274 (24), 17123-17131 (1999)
JOURNAL 99287915
MEDLINE 10358067
PUBMED
REFERENCE 2 (bases 1 to 1739)
AUTHORS Kumar,S. and Zou,C.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109, SmithKline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406, USA

FEATURES
Location/Qualifiers
1. .1739
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="lung"
1. .1739
/gene="Ctgfl"
242. .997
/gene="Ctgfl"
/note="similar to the Mus musculus WISP-2 protein encoded by the sequence presented in GenBank Accession Number AF100778; putative growth factor; Ctgfl-L; contains IGF binding (IGFBP), Von Willebrand Factor type C (VWC) repeat and thrombospondin type I (TSPI) domains; member of the CCN (Ctgf/Cyr61/Nov) family; lacks the fourth carboxy-terminal (CT) domain present in other members of the CCN family"
/codon_start=1
/product="connective tissue growth factor-like protein precursor"

source
gene
CDS

```
/protein_id="AAD18058.1"
/db_xref="GI:4337060"
/translation="MRGNPLHLHLAISFLCILSMVYSQLCPAPCAPWTPPQCPPGVP
LVLDGCGCCRVCCARLGECDHLHVCDPQGLVCQPGAGPSGRGAVCLFEEDDGCEV
NGRRLDGETFKPNCRLVLCRDDGGFTCLPLCSEDVRLPSWDCPRPRRIQVPGRCPE
WVCDQAVMQPAIQPSAQGHLSALVTPASADGPCPNWSTAWGPCSTTCGLGIATRVS
NQNRFCQLEIQRLCLSRPCLASRSHGSWNSAF"
BASE COUNT 375 a 480 c 489 g 395 t
ORIGIN
Alignment Scores:
Pred. No.: 3,98e-67 Length: 1739
Score: 1308.50 Matches: 226
Percent Similarity: 93.63% Conservative: 9
Best Local Similarity: 90.04% Mismatches: 15
Query Match: 90.87% Indels: 1
DB: 10 Gaps: 1
US-10-010-408-2 (1-250) x AF126063 (1-1739)
QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuLeuLeuSerMet 20
Db 242 ATGAGGGGCAACCCACTGATCCATCTTCGGCCATTTCCTTCCTCTGCATCTCTCAATG 301
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
Db 302 GTGTATTCACAGCTGTGCCAGCACCCCTGTGCCTGTCTTGGACACACCCAGTGCCCA 361
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
Db 362 CCGGGGGTACCCCTGGTGTGTGGATGGCTGGCTGCTGTCGAGTGTGCCCCGAGGCTG 421
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 422 GGGGAGTCTCGACCACTGCATGCTCTGGACCCCCAGCCAGGGCCCTGGTTGTCAGCCT 481
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 482 GGGGCAGGCCCCAGTGGCGTGGTGTGTGCCTCTTCGAAGAGGATGACGGGAGCTGT 541
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db 542 GAGGTGAACGGCGCAGGTACCTGGATGGGGAGAGACCTTTAAACCCAATGACGGGTTTG 601
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 602 TGCCCGCTGTGATGACGGTGGTTTCACCTGCCTGCCGTGTGCAGTGAGGATGTGCGGCTG 661
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 662 CCCAGCTGGGACTGCCCCACGCCCCAGGAGAATACAGGTGCCAGGAGGTGCTGCCCCGAG 721
QY 161 TrpValCysAspGlnGlyVal---ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179
Db 722 TGGGTGTGTACCAAGGCAGTGATGCAGCGCGCAATCCAGCCCTCCTCAGCCCCAAGGACAC 781
QY 180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199
Db 782 CAACCTTCTGCCCTTGTCACTCCTGCATCTGCCGATGGCCCCCTGTCCAAACTGGAGCACA 841
QY 200 AlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln 219
Db 842 GCTTGGGGCCCTGCTCAACCACCTGTGGGTGGGCATAGCCACCCGAGTATCCAACCCAG 901
QY 220 AsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
Db 902 AACCGATTCTGCCAACTGGAGATCCAGCGTCGCCTGTGTCTGTCTCCAGACCCCTGCCTGGCA 961
QY 240 AlaArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 962 TCCAGGAGCCACGGCTCATGGAACAGTGCCTTC 994
RESULT 6
```

```
AX076919
LOCUS AX076919 1266 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 31 from Patent WO0105836.
ACCESSION AX076919
VERSION AX076919.1 GI:13121575
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1266)
AUTHORS Botstein,D., Goddard,A., Gurney,A.L., Hillan,K.J., Roy,M.A. and Wood,W.I.
TITLE Polypeptidic compositions and methods for the treatment of tumors
JOURNAL Patent: WO 0105836-A 31 25-JAN-2001;
Genentech, Inc. (US)
FEATURES
source
1..1266
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 216 a 418 c 390 g 242 t
ORIGIN
Alignment Scores:
Pred. No.: 3,56e-53 Length: 1266
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 6 Gaps: 0
US-10-010-408-2 (1-250) x AX076919 (1-1266)
QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuLeuLeuSerMet 20
Db 10 ATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTCTCAAAG 69
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
Db 70 GTGCGTACCCAGCTGTGCCCCGACACCATGTACCTGCCCTGGCCACTCCCCGATGCCCG 129
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
Db 130 CTGGGAGTACCCCTGGTGTGTGGATGGTGTGGCTGTGCCGGGTATGTGCACGGCGCTG 189
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 190 GGGAGCCCTGCGACCAACTCCACGTCTCGGACGCCAGCGCCTGGTCTGCCAGCCC 249
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 250 GGGCAGGACCCGGTGGCCGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCAGCTGT 309
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db 310 GAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGC 369
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 370 TGCCGCTGCGAGGACGGCGCTTACCTGCGTGCCGCTGTGCAGCAGGATGTGCGGCTG 429
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 430 CCCAGCTGGGACTGCCCCACCCACCCAGGAGGGTCTGAGGTCTCTGGGCAAGTGTGCCCTGAG 489
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGCGGCCCAAGGAGGGGGGACTGGGGGACCCAGCCCCCTTCCAGCCCCAAGGACCCAG 549
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCCTTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCCCAGAAATGGAGCACGGCC 609
```


NGRLYREGETFQPHCSIRCRCEDEGGFTCVPLCSEDVRLPSWDCHPHRRVEVLGKCCPE
WVGQGGGLGTQPLPAQGFSGLYVSSLPVGPVPCPEWSTAWGPCSTTCGLGMATRVSN
QNRFCRLETQRRLLCSRPCPPSRGRSPQNSAF"
BASE COUNT 235 a 418 c 389 g 241 t
ORIGIN

Alignment Scores:
Pred. No.: 3.6e-53 Length: 1283
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 9 Gaps: 0

US-10-010-408-2 (1-250) x AF083500 (1-1283)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
|||||:||||| ||| |||||||||
Db 9 ATGAGAGGCACACCGAAGACCCACCTCTGGCCTTCTCCCTCCTCTGCTCTCAAG 68
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
||| ||||||| ||||||||| ||||||| |||||||:|||||
Db 69 GTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCCCTGGCCACTCCCCGATGCCG 128
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgLeu 60
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 129 CTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTGCCGGTATGTGCACGGCGCTG 188
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
||||| ||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 189 GGGGAGCCCTGGGACCAACTCCACGCTCTGGACGCCAGCCAGGGCCTGTCTGCCAGCCC 248
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 249 GGGCAGGACCCGGTGGCCGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCTGT 308
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
||||| ||||||||| ||| :||| ||||||||| |||:||||| |||:|||||
Db 309 GAGGTGAACGCCCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCG 368
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 369 TGCCGCTCGAGGACGGCGGCTTCACTGCGTGCCTGTGTGCAGCGAGGATGTGCGGCTG 428
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
||||| ||||||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 429 CCCAGCTGGGACTGCCCCACCCAGGAGGTCGAGGTCTCTGGCAAGTGTGCCCTGAG 488
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
||||| ||||||| |||||| ||| |||||||||
Db 489 TGGGTGTGCGGCAAGGAGGGGACTGGGACCCAGCCCTTCCAGCCCAAGGACCCAG 548
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
||| |||||||: ||||||||| ||||||||| |||||||||
Db 549 TTTTCTGGCCTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCAGATGGAGCACGGCC 608
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 609 TGGGGACCCCTGCTCGACCACCTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAC 668
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||:|||||
Db 669 CGCTTCTGCCGACTGGAGACCCAGCGCGCTGTGCCTGTCTCCAGGGCCCTGCCACCCCTCC 728
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
||| ||| |||||||||
Db 729 AGGGGTGCGCAGTCCACAAAACAGTGCCTTC 758

RESULT 9
AR210322
LOCUS
DEFINITION Sequence 13 from patent US 6387657.
ACCESSION AR210322

VERSION AR210322.1 GI:21512523
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1293)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I. WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 13 14-MAY-2002;
FEATURES Location/Qualifiers
source 1..1293
BASE COUNT 232 a 425 c 393 g 243 t
ORIGIN
Alignment Scores:
Pred. No.: 3.63e-53 Length: 1293
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 6 Gaps: 0
US-10-010-408-2 (1-250) x AR210322 (1-1293)
QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
|||||:||||| ||| ||||||||| ||| |||||||||
Db 22 ATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCTCTCAAG 81
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||:|||||
Db 82 GTGCGTACCCAGCTGTGCCCCGACACCATGTACCTGCCCTGGCCACCTCCCCGATGCCG 141
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgLeu 60
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 142 CTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTGCCGGTATGTGCACGGGGCTG 201
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
||||| ||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 202 GGGGAGCCCTGGACCAACTCCACGCTGTGCAGCGCCAGCCAGGCGCTGTGCCAGCCC 261
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100
||||| ||||||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 262 GGGCAGGACCCGGTGGCCGGGGCCCTGTGCTTGTGGCAGAGGACGACGAGCTGT 321
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
||||| ||||||||| ||| :||| ||||||||| |||:||||| |||:||||| |||:|||||
Db 322 GAGGTGAACGCCCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCG 381
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 382 TGCCGCTGCAGGACGGCGGCTTCACTGCGTGCCTGTGTGCAGCGAGGATGTGCGGCTG 441
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
||||| ||||||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 442 CCCAGCTGGGACTGCCCCACCCAGGAGGGTCTGAGGTCTTCCAGCAAGTGTGCCCTGAG 501
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
||||| ||||||| |||||| ||| ||||||||| ||| |||||||||
Db 502 TGGGTGTGCGGCCAAGGAGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGGACCCAG 561
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
||| |||||||: ||||||||| ||||||||| ||||||||| ||||||||| |||:|||||
Db 562 TTTTCTGGCCTGTCTCTTCCCTGCCCCCTGGTGTGCCCTGCCAGATGGAGCACGGCC 621
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 622 TGGGGACCCCTGCTCGACCACCTGTGGGCTGGGCATGGCCACCCCGGGTGTCCAACCAAC 681
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||:|||||
Db 682 CGCTTCTGCCGACTGGAGACCCAGCGCGCCTGTGCCTGTGTCCAGGGCCCTGTGCCACCCCTCC 741

QY 241 ArgSerHisSerTrpAsnSerAlaPhe 250
| | | | |
Db 742 AGGGGTGGCAGTCCACAAAACAGTGCCTTC 771
| | | | |
RESULT 10
AR210323/c
LOCUS AR210323 1293 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 14 from patent US 6387657.
ACCESSION AR210323
VERSION AR210323.1 GI:21512524
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1293)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 14 14-MAY-2002;
FEATURES
source
1. .1293
/organism="unknown"
BASE COUNT 243 a 393 c 425 g 232 t
ORIGIN
Alignment Scores:
Pred. No.: 3.63e-53 Length: 1293
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 6 Gaps: 0
US-10-010-408-2 (1-250) x AR210323 (1-1293)
QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
| | | | |
Db 1272 ATGAGAGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTCCTCTCAAAG 1213
| | | | |
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
| | | | |
Db 1212 GTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCG 1153
| | | | |
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60
| | | | |
Db 1152 CTGGGATACCCCTGGTGTGGATGGCTGTGGCTGTGCCGGTATGTGCACGGCGGCTG 1093
| | | | |
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
| | | | |
Db 1092 GGGGAGCCCTGCGACCAACTCCACGCTCGGACGCCAGCCAGCGCCTGGTCTGCCAGCCC 1033
| | | | |
QY 81 GlyAlaGlyProGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
| | | | |
Db 1032 GGGGCAGGACCCGGTGGCCGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCAGCTGT 973
| | | | |
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
| | | | |
Db 972 GAGGTGAACGGCCGCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCCG 913
| | | | |
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
| | | | |
Db 912 TGCCGCTGCGAGGACGGCGGCTTACCTGCTGCCGCTGTGAGCGGAGGATGTGCGGCTG 853
| | | | |
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
| | | | |
Db 852 CCCAGCTGGGACTGCCCCACCCAGGAGGGTCCGAGGTCTTGGGCAAGTGTGCCCTGAG 793
| | | | |
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
| | | | |
Db 792 TGGGTGTGCGGCCAAGGAGGGGGGACTGGGGACCCAGCCCCCTCCAGCCCCAAGGACCCAG 733
| | | | |
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
| | | | |

Db 732 TTTTCTGGCCTTGTCTCTCCCTGCCCTGGTGTCCCTGCCAGAAATGGAGCAGGCC 673
| | | | |
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
| | | | |
Db 672 TGGGGACCCCTGCTCGACACCTGTGGCTGGCATGGCCACCCGGGTGTCCAACCAAC 613
| | | | |
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
| | | | |
Db 612 CGCTTCTGCCGAGTGGAGACCCAGCGCCCTGTGCCTGTCCAGGCCCTGCCACCCCTCC 553
| | | | |
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
| | | | |
Db 552 AGGGGTGGCAGTCCACAAAACAGTGCCTTC 523
| | | | |
RESULT 11
AF074604
LOCUS AF074604 1309 bp mRNA linear PRI 18-JUL-1998
DEFINITION Homo sapiens connective tissue growth factor-related protein precursor (CT58) mRNA, complete cds.
ACCESSION AF074604
VERSION AF074604.1 GI:3328191
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1309)
AUTHORS Rowles,J. and Gendler,S.
TITLE CT58, a new member of the connective tissue growth factor family, interacts with the breast cancer associated mucin MUC1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1309)
AUTHORS Rowles,J. and Gendler,S.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1998) Biochemistry and Molecular Biology, Mayo Clinic Scottsdale, 13400 E. Shea Blvd., Scottsdale, AZ 85259, USA
FEATURES
source
1. .1309
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HeLa"
1. .1309
/gene="CT58"
7. .759
/gene="CT58"
/note="contains three of four modules found in growth regulators related to connective tissue growth factor as described in FEBS Letters 327:125-130,1993; identified in yeast two-hybrid screen using the epithelial mucin MUC1 as bait"
/codon_start=1
/product="connective tissue growth factor-related protein precursor"
/protein_id="AAC26794.1"
/db_xref="GI:3328192"
/translation="MRGTPKTHLLAFSLCLLSKVRTQLCPTPCTCPWPPPPRCPLGVPLVLDGCGCCRVCAARLGEPCDQLHVCDASQGLVCQPGAGPGGRGALCLLAEDDSCEVNGRLYREGETFQPHCSIRCEDGGFTCVPLCSEDVRLPSWDPCPHPRRVEVLGKCPE.WVCGQGGGLGTQPLPAQGPQFSGLVSSLPPGVPCPEWSTAWGPCSTTCGLGMATRVSNQNRFCRLQTRRLCLSRPCPPSRGRSPQNSAF"
BASE COUNT 261 a 418 c 387 g 242 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3.68e-53 Length: 1309
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 9 Gaps: 0
US-10-010-408-2 (1-250) x AF074604 (1-1309)

Db 576 CCCAGCTGGGACTGCCCCACCCAGGAGGTCGAGGTCTGGGCAAGTCTGCCCTGAG 635

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180

Db 636 TGGGTGTGCGGCCCAAGGAGGGGACTGGGGACCCAGCCCTTCCAGGCCAAGGACCCAG 695

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200

Db 696 TTTTCTGGCCTTGTCCTCCCTGCCCCCTGGTGTCCCTGCCAGAGATGGAGCACGGCC 755

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220

Db 756 TGGGACCCCTGCTCGACCACTGTGGGTGGGCATGGCCACCCGGGTGTCCAACCAAGAC 815

QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240

Db 816 CGCTTCTGCGGACTGGAGACCCAGCGCGCCTGTGCCTGTCCAGGCGCCTGCCACCCCTCC 875

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250

Db 876 AGGGTGGCAGTCCACAAACAGTGCCTTC 905

RESULT 13

BC017782

LOCUS

DEFINITION Homo sapiens, WNT1 inducible signaling pathway protein 2, clone

MG:2271 IMAGE:4691574, mRNA, complete cds.

ACCESSION BC017782

VERSION BC017782.1 GI:17389482

KEYWORDS MGC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1450)

Strausberg,R.

Direct Submission

Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: CLONTECH

CDNA Library Preparation: CLONTECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 36 Row: m Column: 3

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507922.

Location/Qualifiers

1. 1450

/organism="Homo sapiens"

/db_xref="LocusID:8839"

/db_xref="taxon:9606"

/clone="MGC:2271 IMAGE:4691574"

/tissue_type="Lung"

/clone_lib="NIH_MGC_77"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

165. .917

/codon_start=1

/product="WNT1 inducible signaling pathway protein 2"

/protein_id="AAH17782.1"

CDS

/db_xref="GI:17389483"

/translation="MRGTPKTHLAFSLCLLSKVRTQLCPTPCTCPWPPRCPLGVP

LVLDCGCCRVCARRLGEPCDQLHVCDSQGLVCPGAGGGRGALCLLAEDDSSCEV

NGRLYREGETFQPHCSIRCEDGGFTCVPLCSEDVRLPSWDGPHPRRVEVLGKCCE

WVCGGGGLGTQPLPAQGPQFSGLVSSLPVGPVPCPEWSTAWGPCSTTCGLGMATRVSN

QNRFCRLTQRRLCLSRPCPPSRGRSPQNSAF"

BASE COUNT 272 a 457 c 441 g 280 t

ORIGIN

Alignment Scores:

Pred. No.: 4.08e-53 Length: 1450

Score: 1064.00 Matches: 184

Percent Similarity: 80.40% Conservative: 17

Best Local Similarity: 73.60% Mismatches: 49

Query Match: 73.89% Indels: 0

DB: 9 Gaps: 0

US-10-010-408-2 (1-250) x BC017782 (1-1450)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20

Db 165 ATGAGAGGCGACACCGAAGACCCACCTCTCGCCTTCTCCCTCCTCTCTCTCAAG 224

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40

Db 225 GTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCG 284

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeu 60

Db 285 CTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTGCCGGGTATGTGCACGGGGCTG 344

QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80

Db 345 GGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCGGCTGGTGTGCCAGCCC 404

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100

Db 405 GGGGAGGACCCGGTGGCGGGGGCCCTGTGCCTTCTTGGCAGAGGACACAGCAGTGT 464

QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120

Db 465 GAGGTGAACGGCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGC 524

QY 121 CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140

Db 525 TGCCGCTGCGAGGACGGCGGCTTACCTCGCTGCCGTGTGCAGCGAGGATGTGCGGCTG 584

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160

Db 585 CCCAGCTGGAGTGTCCCCACCCAGGAGGGTTCAGGTCTCTGGGCAAGTGTGCCCTGAG 644

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180

Db 645 TGGGTGTGCGGGCAAGGAGGGGACTGGGGACCCAGCCCCCTTCCAGCCCAAGGACCCAG 704

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200

Db 705 TTTTCTGGCCTTGTCCTTCCCTGCCCCCTGGTGTCCCTGCCAGAAATGGAGCACGGCC 764

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220

Db 765 TGGGACCCCTGTGACCACTGTGGGTGGGCATGGCCACCCCGGTGTCCAACCAAGAC 824

QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240

Db 825 CGCTTCTGCGGACTGGAGACCCAGCGCGCCTGTGCCTGTCCAGGCGCCTGCCACCCCTCC 884

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250

Db 885 AGGGTGGCAGTCCACAAACAGTGCCTTC 914

RESULT 14

AR210337

LOCUS AR210337 738 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 38 from patent US 6387657.
ACCESSION AR210337
VERSION AR210337.1 GI:21512542
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 738)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 38 14-MAY-2002;
FEATURES Location/Qualifiers
source 1..738
/organism="unknown"
BASE COUNT 104 a 272 c 238 g 124 t
ORIGIN
Alignment Scores:
Pred. NO.: 1.97e-52 Length: 738
Score: 1047.00 Matches: 181
Percent Similarity: 80.08% Conservative: 16
Best Local Similarity: 73.58% Mismatches: 49
Query Match: 72.71% Indels: 0
DB: 6 Gaps: 0
US-10-010-408-2 (1-250) x AR210337 (1-738)
QY 5 ProLeuIleHisLeuLeuAlaThrSerPheLeuLeuLeuSerMetValCysAlaGln 24
Db 1 CCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTCTCAAAGGTGCGTACCCAG 60
QY 25 LeuCysArgThrProCysThrCysProTrpThrProGlnCysProGlnGlyValPro 44
Db 61 CTGTGCCCCGACACCATGTACCTGCCCTGGCCACCTCCCCGATGCCCGCTGGGAGTACCC 120
QY 45 LeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeuGlyGluSerCys 64
Db 121 CTGGTGTGGATGGCTGTGGCTGCTGCCGGGATGTGCACGGCGGCTGGGGAGCCCTGC 180
QY 65 AspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyPro 84
Db 181 GACCAACTCCACGCTCTCGACGCCAGCCAGGCGCTGGTCTGCCAGCCCCGGGCGAGACCC 240
QY 85 GlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluValAsnGly 104
Db 241 GGTGGCCGGGGGCCCTGTGCCTCTTTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGGC 300
QY 105 ArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeuCysArgCysAsp 124
Db 301 CGCCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGCAG 360
QY 125 AspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrpAsp 144
Db 361 GACGGCGGCTTACCTGCGTCCGCTGTGCAGGAGGATGTGCGGCTGCCAGCTGGGAC 420
QY 145 CysProArgProLysArgIleGlnValProGlyLysCysCysProGluTrpValCysAsp 164
Db 421 TGCCCCCACCCCGAGGGGTCGAGGTCTCGAGTCTCGSCAAGTGTGCCCTGAGTGGGTGTCCGGC 480
QY 165 GlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu 184
Db 481 CAAGGAGGGGGACTGGGGACCCAGCCCCCTTCAGGCCCAAGGACCCCAAGTTTCTGGCCTT 540
QY 185 valThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAlaTrpGlyProCys 204
Db 541 GTCTCTTCCCTGCCCTGGTGTCCCTGCCAGATGGAGCACGGCCTGGGGACCCCTGC 600
QY 205 SerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsnArgPheCysGln 224
Db 601 TCGACCACCTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAGAACCGCTTCTGCCGA 660

QY 225 LeuGluIleGlnArgArgLeuLeuCysLeuProArgProCysLeuAlaAlaArgSerHisSer 244
Db 661 CTGGAGACCCAGCGCCCTGTGCTGTCCAGGCCCTGCCACCCCTCCAGGGGTGCGAGT 720
QY 245 SerTrpAsnSerAlaPhe 250
Db 721 CCACAAAACAGTGCCTTC 738
RESULT 15
AR210338
LOCUS AR210338 841 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 39 from patent US 6387657.
ACCESSION AR210338
VERSION AR210338.1 GI:21512543
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 841)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 39 14-MAY-2002;
FEATURES Location/Qualifiers
source 1..841
/organism="unknown"
BASE COUNT 124 a 297 c 280 g 140 t
ORIGIN
Alignment Scores:
Pred. NO.: 2.74e-52 Length: 841
Score: 1045.50 Matches: 184
Percent Similarity: 80.63% Conservative: 20
Best Local Similarity: 72.73% Mismatches: 43
Query Match: 72.60% Indels: 6
DB: 6 Gaps: 2
US-10-010-408-2 (1-250) x AR210338 (1-841)
QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuLeuLeuSerMet 20
Db 12 ATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCTCTCTCAAAG 71
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
Db 72 GTGCGTACCCAGCTGTGCCCCACACCATGTACCTGCCCTGGCCACCTCCCCGATGCCCG 131
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeu 60
Db 132 CTGGGAGTACCCCTGGTGGTGGATGGCTGTGCTGTCCGGGTATGTGCACGGCGGCTG 191
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 192 GGGGAGCCCTGCGACCAACTCCACGCTCTGCAGCGCCAGCCAGGCGCTGGTCTGCCAGCCC 251
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 252 GGGGCAGGACCCCGTGGCCGGGGGCCCTGTGCCTCTTTGGCAGAGGACGACGAGCTGT 311
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db 312 GAGGTGAACGGCCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCCACTGCAGCATCCGC 371
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 372 TGCCGCTGCGAGGACGGCGGCTTACCTGCGTGGCCGCTGTGCAGCGAGGATGTGCGGCTG 431
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 432 CCCAGCTGGGACTGCCCCCAACCCAGGAGGTCGAGGTCTTGGGCAAGTGTGCTGCTGAG 491
QY 161 TrpValCysAspGln-----GlyVal-ThrProAlaIleGlnArgSerThrAlaGlnG 178

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: August 1, 2003, 01:17:19 ; Search time 237 Seconds
(without alignments)
2375.526 Million cell updates/sec

Title: US-10-010-408-2
Perfect score: 1440
Sequence: 1 MRGSPLIHLATSFLLLSM.....LCLPRCLAAARSHSSWNSAF 250

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10010408/runat_25072003_101855_17876/app_query.fasta_1.391
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10010408@cgn_1_1_263@runat_25072003_101855_17876 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1440	100.0	753	20	AAZ07517	Rat HICP polypepti
2	1440	100.0	1708	20	AAZ07516	Rat HICP polypepti
3	1323	91.9	681	20	AAZ07521	Rat HICP mature po
C 4	1308.5	90.9	753	20	AAX76489	Mouse WISP-2 prote
5	1308.5	90.9	1734	20	AAX76488	Mouse WISP-2 prote
C 6	1064	73.9	750	20	AAX76487	Human WISP-2 prote
7	1064	73.9	1257	20	AAZ28435	EGF-like homologue
8	1064	73.9	1266	21	AAA30048	Human PRO261 nucle
9	1064	73.9	1266	22	AAS21403	Human cDNA sequenc
10	1064	73.9	1266	22	AAF60368	PRO261 coding sequ
11	1064	73.9	1266	22	AAC97451	Human angiogenesis
12	1064	73.9	1285	19	AAV29260	Human connective t
13	1064	73.9	1293	20	AAX76486	Human WISP-2 prote
14	1064	73.9	1309	22	AAH28214	Connective tissue
15	1064	73.9	1337	22	AAH46952	Human secreted pro
16	1064	73.9	1352	22	AAH46936	Human secreted pro
17	1064	73.9	1522	20	AAX16595	Human growth facto
18	1051	73.0	1267	21	AAA77566	Human PRO261 cDNA
19	1047	72.7	738	20	AAX76501	Human WISP-2 prote
20	1045.5	72.6	841	20	AAX76502	Human WISP-2 prote
21	564	39.2	2075	16	AAT04226	Connective tissue
22	564	39.2	2075	18	AAT45360	Human connective t
23	564	39.2	2075	18	AAT51234	Connective tissue
24	564	39.2	2075	19	AAV38085	Human connective t
25	564	39.2	2075	20	AAH61317	Human connective t
26	564	39.2	2075	21	AAH11278	Human connective t
27	564	39.2	2075	21	AAH11280	Human connective t
28	564	39.2	2075	22	AAD11223	Human connective t
29	564	39.2	2075	22	AAF59954	Human connective t
30	564	39.2	2075	22	AAC87517	Human connective t
31	564	39.2	2075	24	ABK64553	Human benign prost
32	564	39.2	2312	22	AAH28213	Nucleotide sequenc
33	564	39.2	2312	22	AAH02914	Human shear stress
34	550.5	38.2	2267	18	AAT94700	Murine Fisp12 cDNA
35	550.5	38.2	2267	22	AAD11222	Mouse fibroblast s
36	550.5	38.2	2330	13	AAQ26422	Gene for beta-IG-M
37	550.5	38.2	2345	24	ABK63788	Rat sequence diffe
38	548.5	38.1	2338	20	AAH90030	Rat connective tis
39	542.5	37.7	2350	21	AAH15477	Clone 2-4-7 encodi
40	541.5	37.6	1975	14	AAQ36031	Chicken nov coding
41	532	36.9	4231	24	AAS94798	Human DNA sequence
42	520.5	36.1	684	14	AAQ36032	Chicken nov gene f
43	511	35.5	4212	20	AAV65380	Connective tissue
44	511	35.5	4214	18	AAT59618	Connective tissue
45	511	35.5	4214	18	AAT58534	Human connective t

ALIGNMENTS

RESULT 1	
AAZ07517	
ID	AAZ07517 standard; cDNA; 753 BP.
XX	
AC	AAZ07517;
XX	
DT	26-NOV-1999 (first entry)
XX	
DE	Rat HICP polypeptide coding sequence.
XX	
KW	Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
XX	cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
OS	Rattus sp.
XX	
PN	WO9947556-A2.
XX	
PD	23-SEP-1999.

XX 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Castellot JJ;
XX
DR WPI; 1999-562060/47.
DR P-PSDB; AAY27434.
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications -
XX
PS Claim 5; Fig 1; 108pp; English.
XX
CC The invention provides a rat heparin-induced CCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents the coding sequence of rat HICP.
XX
SQ Sequence 753 BP; 132 A; 235 C; 230 G; 156 T; 0 other;

Alignment Scores:
Pred. No.: 1.07e-77 Length: 753
Score: 1440.00 Matches: 250
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-10-010-408-2 (1-250) x AAZ07517 (1-753)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
DB 1 ATGAGGGGACCCCACTGATCCATCTCTGGCCACTTCTCTCCTCTCTCTCAATG 60

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
DB 61 GTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCACTCCAGTGCCCA 120

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgLeu 60
DB 121 CAGGGGTACCCCTGGTGTGGTGTGGCTGTCTGTAAAGTGTGCACGGAGGCTG 180

QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
DB 181 GGGAGTCTCTGGACCACTGTCATGTCTGGACCCCACTGGGCTGGTTGTCAAGCT 240

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100
DB 241 GGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 300

QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
DB 301 GAGGTGAATGGCCGCGAGGTACCTGGATGGAGAGACCTTTAAACCAATTCAGGGTCTCTG 360

QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
DB 361 TGCCGCTGTGATGACGGTGGCTTACCTGCCTGCCGCTGTGCAGTGAGGATGTCGGCTG 420

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
DB 421 CCCAGCTGGGACTGCCCAAGAGAAATACAGGTGCCAGGAAGTGTGCCCCGAG 480

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180

DB 481 TGGGTATGTACCAAGGAGTGACACCGCGCATCCAGCGCTCCACGCGCAAGACACCAA 540
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
DB 541 CTTTCTGCCCTTGTCACTCCTGCCTCTCTGTATGCTCCTTGTCCAAATGGAGCACAGCC 600
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
DB 601 TGGGGCCCTGCTCAACCACTGTGGCTGGCATAGCCACCCGAGTGTCCAACCAAGAAC 660
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
DB 661 CGATTCTGCCAAAGTGGAGATCCCAACGCGCCTGTGTGTGCCCAGACCTGCCTGGCAGCC 720
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
DB 721 AGGAGCCACAGCTCATGGAACAGTGCTTTC 750

RESULT 2
AAZ07516
ID AAZ07516 standard; cDNA; 1708 BP.
XX AAZ07516;
AC AAZ07516;
XX 26-NOV-1999 (first entry)
XX
DE Rat HICP polypeptide encoding cDNA.
XX
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX
OS Rattus sp.
XX
PN W09947556-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Castellot JJ;
XX
DR WPI; 1999-562060/47.
DR P-PSDB; AAY27434.
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications -
XX
PS Claim 2; Fig 1; 108pp; English.
XX
CC This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein.
CC Agents that stimulate or inhibit HICP protein activity or expression,
CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to
CC modulate cell-associated activity. HICP modulators can be used to treat
CC disorders characterized by aberrant HICP protein activity or expression.
CC Probes capable of hybridizing to HICP mRNA or antibodies specific for
CC HICP can be used to detect HICP activity in a biological sample. HICP
CC can be used to treat disorders, such as a cardiovascular or fibrotic
CC disorder, characterized by aberrant cell proliferation.
XX
SQ Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Alignment Scores:
Pred. No.: 2.54e-77 Length: 1708
Score: 1440.00 Matches: 250
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

```
US-10-010-408-2 (1-250) x AAZ07516 (1-1708)
QY      1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db      249 ATGAGGGGACGCCCACTGATCCATCTTCTGGCCACTTCCTTCCTCTGCCTTCTCTCAATG 308
QY      21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
Db      309 GTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCCCACTGCCCCA 368
QY      41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
Db      369 CAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGTAAGTGTGTGCACGGAGGCTG 428
QY      61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db      429 GGGGAGTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGCGCTGGTTTGTACGCCT 488
QY      81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db      489 GGGGCAGGCCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 548
QY      101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db      549 GAGGTGAATGGCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAATTGCAGGGTCCTG 608
QY      121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db      609 TGGCGCTGTGATGACGGTGGCTTCACCTGCCTGCCGTGTGTCAGTGAGGATGTGCGGCTG 668
QY      141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db      669 CCCAGCTGGGACTGCCACGCCCCCAAGAGAATAACAGGTGCCAGGAAAGTGTGCCCCGAG 728
QY      161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db      729 TGGGTATGTGACCAAGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGGACACCAA 788
QY      181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db      789 CTTTCTGCCCTGTGTCACTCCTGCCTCTGCTGATGCTCCTTGTCCAAATGGAGCACAGCC 848
QY      201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db      849 TGGGGCCCCCTGCTCAACCACTGTGGCTGGGCATAGCCACCCCGAGTGTCCAACCAAGAAC 908
QY      221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db      909 CGATTCTGCCAACTGGAGATCCAAACGCCGCCTGTGTCTGCCCCAGACCCCTGCCTGGCAGCC 968
QY      241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db      969 AGGAGCCACAGCTCATGGAACAGTGCCTTC 998

RESULT 3
AAZ07521
ID      AAZ07521 standard; cDNA; 681 BP.
XX
AC      AAZ07521;
XX
DT      26-NOV-1999 (first entry)
XX
DE      Rat HICP mature polypeptide coding sequence.
XX
KW      Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
XX      cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
OS      Rattus sp.
XX
PN      W09947556-A2.
XX
PD      23-SEP-1999.
```

```
XX      18-MAR-1999; 99WO-US05999.
PF
XX
PR      19-MAR-1998; 98US-0044273.
XX
PA      (TUFT ) TUFTS COLLEGE.
XX
PI      Castellot JJ;
XX
DR      WPI; 1999-562060/47.
DR      P-PSDB; AAY27440.
XX
PT      Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT      used in methods to identify modulators or in diagnostic applications -
XX
PS      Disclosure; Fig 2; 108pp; English.
XX
CC      The invention provides a rat heparin-induced CCN-like protein (HICP)
CC      protein. Agents that stimulate or inhibit HICP protein activity or
CC      expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC      can be used to modulate cell-associated activity. HICP modulators can be
CC      used to treat disorders characterized by aberrant HICP protein activity
CC      or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC      specific for HICP can be used to detect HICP activity in a biological
CC      sample. HICP can be used to treat disorders, such as a cardiovascular or
CC      fibrotic disorder, characterized by aberrant cell proliferation. The
CC      present sequence represents the coding sequence of rat HICP mature
CC      polypeptide.
XX
SQ      Sequence 681 BP; 121 A; 211 C; 215 G; 134 T; 0 other;
```

```
Alignment Scores:
Pred. No.:          9.17e-71          Length:          681
Score:             1323.00           Matches:         227
Percent Similarity: 100.00%           Conservative:    0
Best Local Similarity: 100.00%         Mismatches:     0
Query Match:       91.88%             Indels:         0
DB:                20                 Gaps:          0
```

```
US-10-010-408-2 (1-250) x AAZ07521 (1-681)
QY      24 GlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysProGlnGlyVal 43
Db      1 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCAACCCAGTGCACACAGGGGTA 60
QY      44 ProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeuGlyGluSer 63
Db      61 CCCCTGGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGCACGGAGCTGGGGAGTCC 120
QY      64 CysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGly 83
Db      121 TGCACCACTGCATGTCTGCGACCCAGCCAGGCGCTGGTTTGTACGCTGGGCAGGC 180
QY      84 ProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluValAsn 103
Db      181 CCTGGCGCCATGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT 240
QY      104 GlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeuCysArgCys 123
Db      241 GGCCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATTGCAGGGTCTGTGCCGCTGT 300
QY      124 AspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrp 143
Db      301 GATGACGGTGGCTTCACCTGCCTGCCCTGTGCAGTGAGGATGTGCGGCTGCCAGCTGG 360
QY      144 AspCysProArgProLysArgIleGlnValProGlyLysCysCysProGluTrpValCys 163
Db      361 GACTGCCACACGCCCAAGAGAATAACAGTGCACGAGAAAGTGTGCCCGAGTGGGTATGT 420
QY      164 AspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAla 183
Db      421 GACCAGGGAGTGACACCGGGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCC 480
```


QY	184	LeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAlaTrpGlyPro	203
Db	481	CTTGTCACTCCTGCCTGTGCTGATGCTCCTTGTCCAATTGGAGCACAGCTGGGGCCCC	540
QY	204	CysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsnArgPheCys	223
Db	541	TGCTCAACCACCTGTGGGTGGGCATAGCCACCCGAGTGTCCAACCAACCGATTCGTGC	600
QY	224	GlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAlaArgSerHis	243
Db	601	CAACTGGAGATCCAACGGCGCCTGTGTCTGCCACAGACCCTGCCCTGGCAGCCAGGAGCCAC	660
QY	244	SerSerTrpAsnSerAlaPhe	250
Db	661	AGCTCATGGAACAGTGCCTTC	681

RESULT 4

AA76489/C

ID AAX76489 standard; DNA; 753 BP.

AC AAX76489:

DT 06-AUG-1999 (first entry)

DE Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18.

...	WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW	connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW	leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW	tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW	kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW	connective tissue disorder; catabolic state; inflammation;
KW	testicular-related disorder; angiogenesis; immunological disorder; ss.

OS Mus sp.

PN WO9921998-A1.

PD 06-MAY-1999.

PF 29-OCT-1998; 98WO-US22991.

PR 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

PA (GETH) GENENTECH INC.

PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

DR WPI; 1999-337420/28.

PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

PS Disclosure; Page 179-180; 284pp; English.

The present invention describes Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2 and WISP-3 have homology to connective tissue growth factor (CTGF). Products from the present invention can be used to treat WISP-related disorders such as breast, ovarian, and colon cancer or melanoma. The products can be used to treat arteriosclerosis. The products can also be used to treat other diseases e.g. benign and malignant tumours, leukaemia and lymphoid malignancies, neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal, and blastocoelic disorders, haematopoiesis-related disorders, tissue-growth disorders, skin disorders, desmoplasia, fibrotic lesions, kidney disorders, bone-related disorders such as osteoporosis, trauma such as burns, incisions, and other wounds, connective tissue disorders, catabolic states, testicular-related disorders, and inflammatory, angiogenic and immunologic disorders including arteriosclerosis. The products can also be used for detection and diagnosis especially of

This sequence encodes the EGF-like homologue PRO261.
 The invention relates to antibodies (Ab) that bind to any of the
 polypeptides (I) designated PRO187; PRO533; PRO214; PRO240; PRO211;
 PRO230; PRO261; PRO246 or EBAF-2. The Ab, or other agents that inhibit
 expression and/or activity of (I) are used: (i) to inhibit growth of
 tumours; and (ii) as diagnostic/prognostic reagents for detection or
 quantification of (I) in cells or tissues, by standard immunoassays, with
 overexpression being indicative of cancer. For therapeutic use, the Ab
 may be conjugated to a toxin, chemotherapeutic agent or radioisotope.
 Genes expressing (I), many of which are growth factor homologues, are
 overexpressed in some cases of cancer.

SQ Sequence 1257 BP; 215 A; 416 C; 385 G; 241 T; 0 other;

Alignment Scores:

Pred. No.:	4.86e-55	Length:	1257
Score:	1064.00	Matches:	184
Percent Similarity:	80.40%	Conservative:	17
Best Local Similarity:	73.60%	Mismatches:	49
Query Match:	73.89%	Indels:	0
DB:	20	Gaps:	0

US-10-010-408-2 (1-250) x AAX28435 (1-1257)

Qy	1	MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet	20
Db	1	ATGAGAGGCACACGGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTCTCAAAG	60
Qy	21	ValCysAlaGlnLeuCysArgThrProCysThrCysProTirPThrProProGlnCysPro	40
Db	61	GTGCGTACCCAGCTGCCCCGACACCATGTACCTGCCCTGGCCACCTCCCCGATGCCCG	120
Qy	41	GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu	60
Db	121	CTGGGAGTACCCCTGGTCTGGATGGCTGTGGCTGCTGCCGGGATGTGCACGGGGCTG	180
Qy	61	GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro	80
Db	181	GGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGGCTGGTCTGCCAGCCC	240
Qy	81	GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys	100
Db	241	GGGGCAGGACCCGGTGGCCGGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCAGCTGT	300
Qy	101	GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu	120
Db	301	GAGGTGAACGGCCGCTGTATCGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGC	360
Qy	121	CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu	140
Db	361	TGCCGCTGCGAGGACGGCGGCTTCACCTGCGTGGCGCTGTGCAGCGAGGATGTGGCGCTG	420
Qy	141	ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu	160
Db	421	CCCAGCTGGGACTGCCCCACCCACAGGAGGTTCAGGTCCTGGGCAAGTGTGCCCTGAG	480
Qy	161	TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln	180
Db	481	TGGGTGTGCGGCCAAGAGGGGGGACTGGGGACCCAGGCCCTTCCAGGCCCAAGGACCCCG	540
Qy	181	LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla	200
Db	541	TTTTCTGGCCCTTGTCTCTCCCTGCCCCCTGGTGTCCCTGCCCCAGAAATGGACACGGCC	600
Qy	201	TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn	220
Db	601	TGGGGACCCCTGCTCGACCACCTGTGGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAC	660
Qy	221	ArgPheCysGlnLeuGluIleGlnArgLeuArgLeuProArgProCysLeuAlaAla	240
Db	661	CGCTTCTGCCGACTGGAGACCCAGCGCCGCTGTGCCTGTCTCCAGGGCCTGCCACCCCTCC	720

Db 10 ATGAGAGGCACACCGAAGACCCACCTCCTTGGCCTTCTCCCTCCTCTGCCTCCTCTCAAAG 69

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 70 GTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCG 129

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 130 CTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTG 189

QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 190 GGGGAGCCCTGCGACCACTCCACGTCTGCGAGCGCCAGCCAGCGGCTGTGCTGCCAGCC 249

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 250 GGGGAGCGCGGTGGCGGGGCGCTGTGCTCTTGGCAGAGACGACGACGCTGT 309

QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 310 GAGGTGAACGGCGCCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGC 369

QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 370 TGCCGCTGCGAGGACGGCGGCTTACCTGCGTGCCGCTGTGACGAGGATGTGCGGCTG 429

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 430 CCCAGCTGGGACTGCCCCCACCAGGAGGGTCGAGGTCTTGGCAAGTCTGCCCTGAG 489

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 490 TGGGTGTGCGGCCAAGGAGGGGACTGGGGACCCAGCCCTTCCAGCCCCAAGGACCCAG 549

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 550 TTTTCTGGCCTGTCTCTTCCCTGCCCTGGTGTCCCTGCCAGATGGAGCAGCGCC 609

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 610 TGGGGACCTGCTCGACCACCTGTGGGCTGGCATGGCCACCCCGGTGTCCAACCAAGAAC 669

QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 670 CGCTTCTGCCGACTGGAGACCCAGCGCGCCTGTGCCTGTCCAGGCGCCTGCCACCCCTCC 729

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 730 AGGGTTCGAGTCCACAAAACAGTGCCTTC 759

RESULT 9

AAS21403

ID AAS21403 standard; cDNA; 1266 BP.

XX

AC AAS21403;

XX

DT 24-OCT-2001 (first entry)

XX

DE Human cDNA sequence encoding for PRO261 polypeptide.

XX

KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.

OS Homo sapiens.

XX

PN WO200140466-A2.

XX

PD 07-JUN-2001.

XX

PF 01-DEC-2000; 2000WO-US32678.

XX

PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.

(GETH) GENENTECH INC.

Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

XX WPI; 2001-408281/43.
DR P-PSDB; AAU12331.

XX Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical -

XX Claim 3; Fig 319; 813pp; English.

XX AAS21244-AAS21518 encode for novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIA. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.

XX SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Alignment Scores:

Pred. No.: 4.9e-55 Length: 1266
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0

Db 370 TGCCGCTGCGAGGACGGCGGCTTCACCTGCGTCCGCTGTGCAGGAGGATGTGCGGCTG 429

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
|||||

Db 430 CCCAGCTGGGACTGCCCCACCCAGGAGGGTCGAGGTCTTGGGCAAGTGCTGCCCTGAG 489

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
|||||

Db 490 TGGGTGTGCGGCAAGAGGGGAGTGGGGACCCAGCCCTTCCAGCCCAAGGACCCAC 549

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
||| |||||::

Db 550 TTTTCTGGCCTTGCTCTTCCCTGCCCCCTGGTGTCCCTGCCAGATGGAGCACGGCC 609

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
|||||

Db 610 TGGGGACCTGCTCGACCACCTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAGAAC 669

QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
|||||

Db 670 CGCTTCTGCCGACTGGAGACCCACCGCGCCTGTGCCTGTCCAGGCCCTGCCACCCCTCC 729

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
||| |||||

Db 730 AGGGGTGCGAGTCCACAAACAGTGCCTTC 759

RESULT 11
AAC97451
ID AAC97451 standard; cDNA; 1266 BP.
XX AAC97451;
AC AAC97451;
XX
DT 28-FEB-2001 (first entry)
XX
DE Human angiogenesis-associated protein PRO261 cDNA, SEQ ID NO:107.
XX
KW Human; angiogenesis-associated protein; PRO; endothelial cell growth;
KW cardiac hypertrophy; cardiovascular disorder; osteoporosis; endothelial disorder;
KW angiogenic disorder; atherosclerosis; osteoporosis; hypertension;
KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;
KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
KW gene therapy; transgenic animal; ss.
XX
OS Homo sapiens.
XX
PN WO200053753-A2.
XX
PD 14-SEP-2000.
XX
PF 05-JAN-2000; 2000WO-US00219.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;
PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;

PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;
XX
DR WPI; 2001-090793/10.
P-PSDB; AAB53084.
XX
PT New isolated nucleic acid for producing a PRO polypeptide, analyzing
genetic disorders and treating cardiovascular, endothelial or
PT angiogenic disorders, such as atherosclerosis, wounds or cancer -
XX
PS Claim 58; Fig 41; 293pp; English.
XX
CC The invention relates to novel human angiogenesis-associated proteins
CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding
CC PRO proteins. The invention also relates to vectors and host cells
CC comprising a PRO nucleic acid, the recombinant production of a PRO
CC protein, PRO antibodies specific for a PRO protein, fusion proteins
CC comprising a PRO protein, agonists or antagonists of a PRO protein, and
CC compounds which inhibit the expression of a PRO gene. The invention
CC additionally encompasses methods of identifying modulators of PRO
CC expression or activity; diagnosing a cardiovascular, endothelial or
CC angiogenic disorder, or a susceptibility to such a disorder by detecting
CC mutations in a PRO gene, or the expression level of a PRO gene within a
CC particular tissue; treating a cardiovascular, endothelial or angiogenic
CC disorder via the administration of a PRO protein, PRO nucleic acid, or
CC PRO agonist or antagonist; a retroviral gene therapy vector comprising a
CC PRO nucleic acid; and methods of inhibiting or stimulating endothelial
CC cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the
CC administration of a PRO protein, or an agonist or antagonist thereof.
CC PRO nucleic acids, PRO proteins, antibodies against PRO proteins, PRO
CC agonists and PRO antagonists may be used as therapeutic agents to treat
CC cardiovascular, endothelial or angiogenic disorders, such as
CC atherosclerosis, osteoporosis, myocardial infarction, hypertension,
CC diabetic retinopathy, rheumatoid arthritis, Crohn's disease, psoriasis,
CC endometriosis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's
CC disease, or stroke. PRO nucleic acids are additionally useful in the
CC recombinant production of PRO proteins, as hybridisation probes to
CC screen libraries to isolate cDNAs with sequence identity to PRO proteins,
CC to map genes encoding PRO proteins, to analyse genetic disorders, and in
CC gene therapy. PRO nucleic acids can also be used to produce transgenic
CC animals useful for the development and screening of potential
CC therapeutic agents. The present sequence represents a cDNA encoding a PRO
CC protein of the invention.
XX
SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Alignment Scores:
Pred. No.: 4.9e-55 Length: 1266
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 22 Gaps: 0

US-10-010-408-2 (1-250) x AAC97451 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
|||||

Db 10 ATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTCTCTCAAG 69

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrThrProGlnCysPro 40
||| |||||

Db 70 GTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGCCACCTCCCCGATGCCCG 129

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
|||||

Db 130 CTGGGAGTACCCCTGGTGTGGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTG 189

QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
|||||

Db 190 GGGAGCCCTGGGACCAACTCCACGTCTGCGACGCCAGCCAGGCCCTGGTCTGCCAGCCC 249

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlySerCys 100
|||||

Db 250 GGGCAGGACCCGGTGGCCGGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCAGCTGT 309

QY 101 GluValAsnGlyArgTyrLeuAspGlyGluThrPheLeysProAsnCysArgValLeu 120
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 310 GAGGTGAACGGCCGCTGTATCGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGC 369

QY 121 CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 370 TGCCGCTGCGAGGACGGCGCTTCACTGCTGCGCTGTGCAGCGAGGATGTGCGGCTG 429

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 430 CCCAGCTGGGACTGCCCCACCCAGGAGGGTCGAGGTCCTGGGCAAGTGTGCCCTGAG 489

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 490 TGGGTGTGCGGCCCAAGAGGGGACTGGGGACCCAGCCCTTCCAGCCCCAAGGACCCAG 549

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 550 TTTTCTGGCCTTGCTCTTCCCTGCCCCCTGGTGTCCCTGCCCCAGAATGGAGCACGGCC 609

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 610 TGGGGACCCCTGCTCGACCACCTGTGGGCTGGGATGGCCACCCGGGTGTCCAACCAAGAAC 669

QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 670 CGCTTCTGCCGACTGGAGACCCAGCGCCGCTGTGCCTGTCCAGGCCCTGCCACCCCTCC 729

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 730 AGGGGTGCAGTCCACAAACAGTGCCTTC 759

RESULT 12

AAV29260

ID AAV29260 standard; DNA; 1285 BP.

XX AAV29260;

XX 14-SEP-1998 (first entry)

DE Human connective tissue growth factor-3 gene.

KW Connective tissue growth factor-3; CTGF-3; human; cancer;

KW arthritis; fibrosis; osteoporosis; diagnosis; therapy; ds.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 9..761

FT /*tag= a

FT sig_peptide 9..65

FT /*tag= b

FT mat_peptide 66..758

FT /*tag= c

XX WO9821236-A1.

XX 22-MAY-1998.

PF 08-NOV-1996; 96WO-US17856.

XX 08-NOV-1996; 96WO-US17856.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Chopra A, Ebner R, Ruben SM;

DR WPI; 1998-297864/26.

DR P-PSDB; AAW37946.

XX Novel human connective tissue growth factor 3 gene - useful for the

PT diagnosis and treatment of e.g. cancer, arthritis, fibrosis, osteoporosis

XX Claim 2; Fig 1A-B; 87pp; English.

CC This nucleotide sequence codes for human connective tissue growth factor-3 (CTGF-3) protein (see AAW37946), a novel member of the growth factor superfamily. It was discovered in a cDNA library derived from human osteoblasts. The gene has also been identified in cDNA libraries from ovary, testis, heart, lung, skeletal muscle, adrenal medulla, adrenal cortex, thymus, prostate, small intestine and colon. A cDNA clone is deposited as ATCC 97756. Also provided are vectors, host cells and recombinant methods for producing CTGF-3 polypeptides. CTGF-3 nucleic acid sequences or their fragments, e.g. primers or probes, can be used to diagnose diseases where CTGF-3 expression is enhanced, e.g. cancer, arthritis, fibrosis or atherosclerosis, or diseases where expression is decreased such as in osteoporosis. Disorders characterised by decreased or increased levels of CTGF-3 can be treated by administering CTGF-3 polypeptides and anti-CTGF-3 antibodies, respectively.

XX SQ Sequence 1285 BP; 237 A; 418 C; 389 G; 241 T; 0 other;

Alignment Scores:

Pred. No.: 4.98e-55 Length: 1285

Score: 1064.00 Matches: 184

Percent Similarity: 80.40% Conservative: 17

Best Local Similarity: 73.60% Mismatches: 49

Query Match: 73.89% Indels: 0

DB: 19 Gaps: 0

US-10-010-408-2 (1-250) x AAV29260 (1-1285)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 9 ATGAGAGGCACACCCGAAGACCCACCTCCTGGCCCTTCTCCCTCCTCTCTCTCAAG 68

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 69 GTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCG 128

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 129 CTGGGAGTACCCCTGGTGTGGATGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTG 188

QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 189 GGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGCCAGCCC 248

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 249 GGGGCAGGACCCGTTGCCCGGGGGCCCTGTGCCCTTTGGCAGAGGACGACGAGCTGT 308

QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 309 GAGGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGC 368

QY 121 CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 369 TGCCGCTGCGAGGACGGCGGCTTCACTGCGTGGCTGTGCAGCGAGGATGTGCGGCTG 428

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 429 CCCAGCTGGGACTGCCCCACCCAGGAGGGTCGAGGTCCTGGGCAAGTGTGCCCTGAG 488

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 489 TGGGTGTGCGGCCCAAGAGGGGACTGGGGACCCAGCCCTTCCAGCCCCAAGGACCCAG 548

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
111 :::||||| 111 :::||||| 111 :::||||| 111 :::||||| 111 :::

Db 549 TTTTCTGGCCTTGCTCTTCCCTGCCCCCTGGTGTCCCTGCCCCAGAATGGAGCACGGCC 608

Qy 201 trpGlyProCysSerThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
|||||
Db 609 TGGGACCCTGCTCGACCACCTGTGGCTGGCATGGCCACCGGGTGTCCAACCAAGAC 668

Qy 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
|||||
Db 669 CGCTTCTGCGGACTGGAGACCCAGCGCGCTGTGCCTGTCCAGGCCCTGCCACCCCTCC 728

Qy 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
|||
Db 729 AGGGTGGCAGTCCACAAACACAGTGCCTTC 758

RESULT 13
AAX76486

ID AAX76486 standard; DNA; 1293 BP.

AC AAX76486;

DT 06-AUG-1999 (first entry)

DE Human WISP-2 protein nucleotide sequence SEQ ID NO:13.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.

OS Homo sapiens.

XX WO9921998-A1.

PN 06-MAY-1999.

XX 29-OCT-1998; 98WO-US22991.

PR 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

XX (GETH) GENENTECH INC.

XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

XX WPI; 1999-337420/28.

DR P-PSDB; AAY17649.

XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2, and 3

PS Example 4; Page 174-175; 284pp; English.

XX The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.

CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ. Sequence 1293 BP; 232 A; 425 C; 393 G; 243 T; 0 other;

Alignment Scores:

Pred. No.: 5.01e-55 Length: 1293
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 20 Gaps: 0

US-10-010-408-2 (1-250) x AAX76486 (1-1293)

Qy 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20
|||||
Db 22 ATGAGAGGCACAGCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTCCTCAAG 81

Qy 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
|||
Db 82 GTGCGTACCCAGGTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCG 141

Qy 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeu 60
|||||
Db 142 CTGGAGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 201

Qy 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
|||||
Db 202 GGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCGGCTGGTCTGCCAGCCC 261

Qy 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100
|||||
Db 262 GGGCAGGACCCGCTGGCCGGGGGCCCTGTGCCTTTGGCAGGAGACGACAGCTGT 321

Qy 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
|||||
Db 322 GAGGTGAACGGCGCCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCG 381

Qy 121 CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
|||||
Db 382 TGCCGCTGCGAGGACGGCGGCTTCACCTGCGTGGTGGTGGTGGTGGTGGTGGTGG 441

Qy 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
|||||
Db 442 CCCAGCTGGGACTGCCCCCACCAGGAGGGTTCGAGGTCTCTGGCAAGTCTGCCCTGAG 501

Qy 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
|||||
Db 502 TGGGTGTGGGCCCAAGGAGGGGGAGTGGGGACCCAGCCCTTCCAGCCCCAAGGACCC 561

Qy 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
|||
Db 562 TTTTCTGGCCTGTGTCTCTTCCCTGCCCTTGGTGTGGTGTGGTGTGGTGTGGTGTGG 621

Qy 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
|||||
Db 622 TGGGACCCCTGCTCGACCACTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 681

Qy 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
|||||
Db 682 CGCTTCTGCGGACTGGAGACCCAGCGCGCCTGTGCCTGTGCCTGTGCCTGTGCCTGTGC 741

Qy 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
|||
Db 742 AGGGTGGCAGTCCACAAACACAGTGCCTTC 771

RESULT 14

AAH28214

ID AAH28214 standard; cDNA; 1309 BP.

XX

AC AAH28214;

XX

DT 05-SEP-2001 (first entry)
XX
DE Connective tissue derived growth factor related protein cDNA.
XX
KW Growth factor; protein inhibitor; protease; damaged tissue;
KW platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
KW connective tissue derived growth factor; CTGF; chrysalin; VEGF;
KW keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
KW transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
KW granulocyte macrophage colony stimulating factor; GM-CSF; UPA;
KW vascular endothelial growth factor; urokinase plasminogen activator;
KW dermal ulcer; wound; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 7..759
FT /*tag= a
FT /product= "connective tissue derived growth factor
FT related protein"
XX
PN WO200149309-A2.
XX
PD 12-JUL-2001.
XX
PF 21-DEC-2000; 2000WO-IB01935.
XX
PR 29-DEC-1999; 99GB-0030768.
XX
PA (PFIZ) PFIZER LTD.
PA (PFIZ) PFIZER INC.
XX
PI Davies MJ, Huggins JP, McIntosh FS, Occleston NL;
XX
DR WPI; 2001-418351/44.
DR P-PSDB; AAB84599.
XX
PT Composition for the treatment of damaged tissue i.e. chronic wounds and
PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
PT factor -
XX
PS Disclosure; Page 546; 572pp; English.
XX
CC The specification describes a pharmaceutical composition, comprising
CC a growth factor, an inhibitor agent, i.e. a protease. The inhibitor
CC agent inhibits the action of at least one specific adverse protein,
CC i.e. a protease, that is upregulated in a damaged tissue such as a
CC wound environment. Growth factors which are included in the composition
CC of the invention are platelet-derived growth factor (PDGF), fibroblast
CC growth factor (FGF), connective tissue derived growth factor (CTGF),
CC keratinocyte-derived growth factor (KGF), transforming growth
CC factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor
CC (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth
CC factor (VEGF), and chrysalin. Inhibitors which are included in the
CC composition of the invention include inhibitors of urokinase-type
CC plasminogen activator (uPA) and matrix metalloproteinase (MMP). The
CC composition is useful for the treatment of chronic damaged tissue, i.e.
CC wounds and dermal ulcers. The present sequence encodes a human
CC CTGF-like protein, and is used to produce the composition of the
CC invention.
XX
SQ Sequence 1309 BP; 261 A; 418 C; 387 G; 242 T; 1 other;

Alignment Scores:
Pred. No.: 5.07e-55 Length: 1309
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 22 Gaps: 0

US-10-010-408-2 (1-250) x AAH28214 (1-1309)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
DB 7 ATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCTCCTCCTCCTCAAAG 66
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
DB 67 GTGGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCG 126
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
DB 127 CTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTGCCGGGTATGTGCACGGCGCTG 186
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
DB 187 GGGAGCCCTGCGACCACTCCACGTCTCGACGCCAGCGGCCTGGTCTGCCAGCCC 246
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100
DB 247 GGGCAGGACCCGGTGGCGGGGGCCCTGTGCCTCTTGGCAGAGGACGACGAGCTGT 306
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
DB 307 GAGGTGAACGGCCGCTGTATCGGGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGC 366
QY 121 CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
DB 367 TGGCGCTGCGAGGACGGCGGCTTACCTCGCTGCGCTGCGAGGAGGATGTGCGGCTG 426
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
DB 427 CCCAGCTGGGACTGCCCCGCCAGGAGGGTGGAGGCTGGGCAAGTGTGCCCTGAG 486
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
DB 487 TGGGTGTGGGCCCAAGGAGGGGGGACTGGGGACCCAGCCCCCTCCAGCCCCAGCCAG 546
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
DB 547 TTTTCTGGCCTTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCAGAAATGGAGCAGGCC 606
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
DB 607 TGGGACCCCTGCTCGACCACTGTGGGCTGGGCATGGCCACCGGGGTGTCCAACAGAAC 666
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
DB 667 CGCTTCTGCCGACTGGAGACCCAGCGCGCTGTGCCTGTCCAGGCCCTGCCACCCCTCC 726
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
DB 727 AGGGGTGCGCAGTCCACAAACACAGTGCCTTC 756
RESULT 15
AAH46952
ID AAH46952 standard; cDNA; 1337 BP.
XX
AC AAH46952;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human secreted protein encoding cDNA (clone Id HB0DE48).
KW Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW nootropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW ophthalmological; gene therapy; ss.
OS Homo sapiens.
XX
PN WO200155430-A1.
XX
PD 02-AUG-2001.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2003, 02:52:19 ; Search time 53 Seconds
(without alignments)
1446.589 Million cell updates/sec

Title: US-10-010-408-2
Perfect score: 1440
Sequence: 1 MRGSPILHLATSLFLISM.....ICLPRLCLARSHSSWNSAF 250

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10010408/rnat_25072003_101856_17954/app_query.fasta_1.391
-DB=Issued_Patents_NA -QEXT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10010408_GCGN_1_1_40_@runat_25072003_101856_17954 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1308.5	90.9	1734	4	US-09-182-145-17	Sequence 17, Appl
c 2	1308.5	90.9	1734	4	US-09-182-145-18	Sequence 18, Appl
3	1064	73.9	1293	4	US-09-182-145-13	Sequence 13, Appl
c 4	1064	73.9	1293	4	US-09-182-145-14	Sequence 14, Appl
5	1047	72.7	738	4	US-09-182-145-38	Sequence 38, Appl
6	1045.5	72.6	841	4	US-09-182-145-39	Sequence 39, Appl
7	564	39.2	2075	1	US-08-167-628-1	Sequence 1, Appl
8	564	39.2	2075	1	US-08-386-680-1	Sequence 1, Appl
9	564	39.2	2075	1	US-08-459-717-1	Sequence 1, Appl
10	564	39.2	2075	1	US-08-712-302-1	Sequence 1, Appl
11	564	39.2	2075	2	US-08-880-031-1	Sequence 1, Appl
12	564	39.2	2075	3	US-09-097-179-1	Sequence 1, Appl

13	564	39.2	2075	4	US-09-080-715-1	Sequence 1, Appl
14	564	39.2	2075	4	US-09-142-569-7	Sequence 7, Appl
15	564	39.2	2075	5	PCT-US96-08140-1	Sequence 1, Appl
16	564	39.2	2998	3	US-09-054-368-1	Sequence 1, Appl
17	564	39.2	2998	3	US-09-054-274-1	Sequence 1, Appl
18	564	39.2	2998	4	US-09-056-704-1	Sequence 1, Appl
19	550.5	38.2	2267	4	US-09-142-569-5	Sequence 5, Appl
20	542.5	37.7	2350	4	US-09-187-478-1	Sequence 1, Appl
21	542.5	37.7	2350	4	US-09-292-036-1	Sequence 1, Appl
22	511	35.5	4214	4	US-09-122-135-1	Sequence 1, Appl
23	503.5	35.0	1766	4	US-09-182-145-9	Sequence 9, Appl
c 24	503.5	35.0	1766	4	US-09-182-145-10	Sequence 10, Appl
25	501.5	34.8	1403	4	US-09-182-145-23	Sequence 23, Appl
26	500.5	34.8	1480	4	US-09-142-569-1	Sequence 1, Appl
27	498.5	34.6	2830	4	US-09-182-145-1	Sequence 1, Appl
c 28	498.5	34.6	2830	4	US-09-182-145-2	Sequence 2, Appl
29	496.5	34.5	1418	4	US-09-142-569-3	Sequence 3, Appl
c 30	465	32.3	1128	2	US-08-459-101A-1	Sequence 1, Appl
31	446	31.0	1101	4	US-09-182-145-29	Sequence 29, Appl
32	377.5	26.2	1335	4	US-09-182-145-30	Sequence 30, Appl
c 33	377.5	26.2	1335	4	US-09-182-145-31	Sequence 31, Appl
34	372.5	25.9	1212	4	US-09-182-145-34	Sequence 34, Appl
c 35	372.5	25.9	1212	4	US-09-182-145-35	Sequence 35, Appl
36	371.5	25.8	1142	4	US-09-253-316-1	Sequence 1, Appl
37	326.5	22.7	1062	4	US-09-253-316-3	Sequence 3, Appl
38	220	15.3	693	4	US-09-182-145-24	Sequence 24, Appl
39	220	15.3	1202	4	US-09-182-145-26	Sequence 26, Appl
40	201.5	14.0	2541	2	US-08-656-393-1	Sequence 1, Appl
41	150	10.4	683	4	US-09-182-145-25	Sequence 25, Appl
c 42	150	10.4	1183	4	US-09-182-145-27	Sequence 27, Appl
43	139.5	9.7	8257	4	US-09-484-970B-65	Sequence 65, Appl
44	139	9.7	1271	1	US-08-464-339A-1	Sequence 1, Appl
45	139	9.7	1271	5	PCT-US94-14388-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-182-145-17

Sequence 17, Application US/09182145B

Patent No. 6387657

GENERAL INFORMATION:

APPLICANT: Botstein, David A.
APPLICANT: Cohen, Robert
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Lawrence, David A.
APPLICANT: Levine, Arnold J.
APPLICANT: Pennica, Diane
APPLICANT: Roy, Margaret Ann
APPLICANT: Wood, William I.
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: P1176R2
CURRENT APPLICATION NUMBER: US/09/182,145B
CURRENT FILING DATE: 1998-10-29
EARLIER APPLICATION NUMBER: US 60/063,704
EARLIER FILING DATE: 1997-10-29
EARLIER APPLICATION NUMBER: US 60/073,612
EARLIER FILING DATE: 1998-02-04
EARLIER APPLICATION NUMBER: US 60/081,695
EARLIER FILING DATE: 1998-04-14
NUMBER OF SEQ ID NOS: 156
SEQ ID NO 17
LENGTH: 1734
TYPE: DNA
ORGANISM: Mus musculus
US-09-182-145-17

Alignment Scores:

Pred. No.: 6e-95
Score: 1308.50
Percent Similarity: 93.63%
Length: 1734
Matches: 226
Conservative: 9

Best Local Similarity: 90.04% Mismatches: 15
Query Match: 90.87% Indels: 1
DB: 4 Gaps: 1

US-10-010-408-2 (1-250) x US-09-182-145-17 (1-1734)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20
DB 257 ATGAGGGGCAACCCACTGATCCATCTTGTGGCCATTTCCTTCCCTGCATCTCTCAATG 316
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
DB 317 GTGTATTCCAGCTGTGCCAGCACCCTGTGCTGCTTGCCTTGACACACCCAGTGCCCA 376
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
DB 377 CCGGGGTACCCCTGCTGCTGATGGCTGTGGCTGTGTCGAGTGTGTGCACGGAGGCTG 436
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
DB 437 GGGGAGTCTCGCAGCACCACTGCATGTCTGCCAGACCCAGAGGGCTGTGTGTCAGCCT 496
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
DB 497 GGGGCGAGGCCCCAGTGCCGCTGTGTGTGTGCTCTTCCGAAGAGGATGACGGGAGCTGT 556
QY 101 GluValaGlnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
DB 557 GAGGTGAATGGCCCGCAGGTACCTGGATGGGAGACCTTTAAACCAATTGCAGGGTTTG 616
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
DB 617 TGCCGCTGTGATGACGGGTGTTCACCTGCTGCTGCTGCTGTCAGTGAAGATGCGGCTG 676
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
DB 677 CCCAGCTGGACTGCCCCACGCCCCAGAGAATACAGGTGCCAGGAAGTGCTGCCCGAG 736
QY 161 TrpValCysAspGlnGlyVal--ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179
DB 737 TGGGTGTGTGACCAAGCAGTGAATGACGGGCAATCCAGCCCTCTCAGCCCAAGACAC 796
QY 180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199
DB 797 CAACTTTCCTCCCTTGTCACTCCTGCACTGCTGCCGATGGCCCCCTGTCCAAACTGGAACACA 856
QY 200 AlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln 219
DB 857 GCCTGGGGCCCCCTGCTCAACCACTGTGGGTTGGGCATAGCCACCCGAGTATCCAAACAG 916
QY 220 AsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
DB 917 AACCGATTCTGCCAACTGGAGATCCAGCTGCTGTGTCTGTCCAGAACCCCTGCGCTGGA 976
QY 240 AlaArgSerHisSerSerTrpAsnSerAlaPhe 250
DB 977 TCCAGGAGCCACGGCTCATGGAACAGTGCCCTTC 1009

RESULT 2

US-09-182-145-18/c
Sequence 18, Application US/09182145B
Patent No. 6387657

GENERAL INFORMATION:

APPLICANT: Botstein, David A.
APPLICANT: Cohen, Robert
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Lawrence, David A.
APPLICANT: Levine, Arnold J.
APPLICANT: Pennica, Diane
APPLICANT: Roy, Margaret Ann
APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182.145B
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 18
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-182-145-18

Alignment Scores:

Pred. No.: 6e-95 Length: 1734
Score: 1308.50 Matches: 226
Percent Similarity: 93.63% Conservative: 9
Best Local Similarity: 90.04% Mismatches: 15
Query Match: 90.87% Indels: 1
DB: 4 Gaps: 1

US-10-010-408-2 (1-250) x US-09-182-145-18 (1-1734)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20
DB 1478 ATGAGGGGCAACCCACTGATCCATCTTGTGGCCATTTCCTTCTCTGCAATTCTCTCAATG 1419
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
DB 1418 GTGTATTCCAGCTGTGCCAGCACCCTGTGCTGCTGCTGTCAGTGAAGATGACGGAGGCTG 1359
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
DB 1358 CCGGGGTACCCCTGCTGCTGATGGCTGTGGCTGTGCTGCTGTCAGTGTGTGCACGGAGGCTG 1299
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
DB 1298 GGGGAGTCTCGCAGCACCACTGCATGTCTGCCAGACCCAGCAGGGCTGTGTGTCAGCCT 1239
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
DB 1238 GGGGCGAGGCCCCAGTGCCGCTGTGTGTGTGCTCTTCCGAAGAGATGACGGAGCTGT 1179
QY 101 GluValaGlnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
DB 1178 GAGGTGAATGGCCCGCAGGTACCTGATGGGAGAGACCTTTAAACCAATTGCAGGGTTTG 1119
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
DB 1118 TGCCGCTGTGATGACGGGTGTTCACCTGCTGCCGCTGTGCAATGAGGATGTGCGGCTG 1059
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
DB 1058 CCCAGCTGGACTGCCACGCCCCAGAGAATACAGGTGCCAGGAAGTGCTGCCCGGAG 999
QY 161 TrpValCysAspGlnGlyVal--ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179
DB 998 TGGGTGTGTGACCAAGCAGTGAATGACCGGCAATCCAGCCCTCTCAGCCCAAGACAC 939
QY 180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199
DB 938 CAACTTTCCTCCCTTGTCACTCCTGCACTGCTGCCGATGGCCCCCTGTCCAAACTGGAACACA 879
QY 200 AlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln 219
DB 878 GCCTGGGGCCCCCTGCTCAACCACTGTGGGTTGGGCATAGCCACCCGAGTATCCAAACAG 819
QY 220 AsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
DB 1118 TGGGTGTGTGACCAAGCAGTGAATGACCGGCAATCCAGCCCTCTCAGCCCAAGACAC 819

```
Db      818 AACCATCTGCCAACTGGAGATCCAGCGTCGCTGTGTGTCTCCAGACCCCTGCCTGGCA 759
QY      240 AlaArgSerHisSerTrpAsnSerAlaPhe 250
       :::|||||  |||||||
Db      758 TCACGAGGCCACGCGCTCATGAGACGTGCCTTC 726

RESULT 3
US-09-182-145-13
: Sequence 13, Application US/09182145B
: Patent No. 6387657
: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182,145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 13
: LENGTH: 1293
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-182-145-13

Alignment Scores:
Pred. No.:          9.63e-76          length:          1293
Score:             1064.00           Matches:          184
Percent Similarity: 80.40%           Conservative:      17
Best Local Similarity: 73.60%         Mismatches:       49
Query Match:       73.89%            Indels:           0
DB:                4                Gaps:              0

US-10-010-408-2 (1-250) x US-09-182-145-13 (1-1293)
QY      1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet 20
       |||||||:::|  |||||||  |||  |||||||
Db      22 ATGAGAGGCACACCGAAGACCACTCTCGCCTTCTCTCCTCTCTGCTCTCAAG 81
QY      21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
       |||  |||||||  |||||||  |||||||  |||||||  |||:::|
Db      82 GTGGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCG 141
QY      41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60
       |||||||  |||||||  |||||||  |||||||  :::|
Db      142 CTGGGAGTACCCCTGTGTGTGATGGCTGTGGCTGTGCCGGGTATGTGACGCGGCTG 201
QY      61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
       |||||  |||||  |||||||  |||||||  |||||||  |||||
Db      202 GGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCCAGGCCCTGTCTGCCAGCCC 261
QY      81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
       |||||||  |||||||  |||||||  |||||||  |||||||  |||||
Db      262 GGGGACGAGCCCGGTGGCGGGGGGCCCTGTGCTCTTGGCAGAGGACGACGACGACTGT 321
QY      101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
       |||||||  |||||  :::|  |||||||  |||||||  :::|
Db      322 GAGGTGAACGGCGCGCTGTATCGGGGAAGGGAGAGACCTTCACGCCCTGCAGCATCCGC 381
QY      121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
```

```
Db      382 TGCCGCTGCGAGAGACGGCGCTTCACCTGCGGTGCCGCTGTGCAGCAGGATGTGGGCTG 441
QY      141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
       |||||||  |||||||  |||:::|  |||||||  |||||||
Db      442 CCCAGCTGGGACTGCCCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 501
QY      161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
       |||||||  |||||  |||
Db      502 TGGTGTGCGGCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 561
QY      181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
       |||  |||:::|
Db      562 TTTTCTGGCCTTGTCTTCCCTGCCCCCTGTGTGTCCTGCCCCAGATGAGACAGGCC 621
QY      201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
       |||||||  |||||||  |||||||  |||||||  |||||||
Db      622 TGGGAGCCCTGCTCGACCACTGTGGGCTGGGATGGCCAGCCCGGGGTGCCAACACGAC 681
QY      221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
       |||||||  |||  |||||||
Db      682 CGCTTCTGCCGACTGGAGACCAAGCGCCGCTGTGCTGTCCAGGAGGCCCTGACCCTCC 741
```

```
RESULT 4
US-09-182-145-14/C
: Sequence 14, Application US/09182145B
: Patent No. 6387657
: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182,145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 14
: LENGTH: 1293
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-182-145-14

Alignment Scores:
Pred. No.:          9.63e-76          length:          1293
Score:             1064.00           Matches:          184
Percent Similarity: 80.40%           Conservative:      17
Best Local Similarity: 73.60%         Mismatches:       49
Query Match:       73.89%            Indels:           0
DB:                4                Gaps:              0

US-10-010-408-2 (1-250) x US-09-182-145-14 (1-1293)
```

```
QY      1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet 20
       |||||||  |||||  |||||||  |||  |||||||
Db      1272 ATGAGAGGCACACCGAAGACCACTCTGCGCTTCTCCCTCTCTGCTCTCTCAAG 1213
```


GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182,145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 39
: LENGTH: 841
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1-841
: OTHER INFORMATION: Sequence is synthesized.
: Patent No. 6387657
: US-09-182-145-39

Alignment Scores:
Pred. No.: 1.6e-74 Length: 841
Score: 1045.50 Matches: 184
Percent Similarity: 80.63% Conservative: 20
Best Local Similarity: 72.73% Mismatches: 43
Query Match: 72.60% Indels: 6
DB: 4 Gaps: 2

US-10-010-408-2 (1-250) x US-09-182-145-39 (1-841)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20
DB 12 ATGAGAGGACACCGAAGACCCACCTCTGCGCTTCTCCCTCTGCTCTCAAG 71
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
DB 72 GTGCGTACCAGCTGTGCCCCGACACCAATGTAACCTGCCCTGGCCACCTCCCGATGCCG 131
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
DB 132 CTGGAGTACCCCTGTGTGTGTGATGGCTGTGGCTGCTGCCGGGTATGTGCACGGGCTG 191
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
DB 192 GGGGAGCCCTGCGACCACTCCACGTTCTGCGACGCCAGCCAGGGCTGGTCTGCCAGCCC 251
QY 81 GlyAlaGlyProGlnGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
DB 252 GGGGCAGGACCCGGTGGCCGGGGGGCCCTGTGCTTGTGGCAGAGGACGACAGCAGCTGT 311
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
DB 312 GAGGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGCATCCGC 371
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
DB 372 TGCCGCTGCGAGGACGCGGCTTACCTGCGCTGCCGCTGTGCAGGAGATGTGCGGCTG 431
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160

DB 432 CCCAGCTGGAGCTGCCCCACCCACCCAGAGGGTCGAGGTCTGGGCAAGTCTGCCCTGAG 491
QY 161 TrpValCysAspGln-----GlyVal-ThrProAlaIleGlnArgSerThrAlaGlnG1 178
DB 492 TGGGTGTGGGCGCAAGAGAGGGGAGTGGGACACGACCCCT-----CCAGCCCAAG 542
QY 178 YHisGlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSe 198
DB 543 ACCCAGTTTCTGGCCCTGTCTCTCTCCCTGCCCCCTGGTGTGCCCTGCCAGATGGAG 602
QY 198 rThrAlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAs 218
DB 603 CACGGCCTGGGGAACCTGTGCTGACCACTGTGGCTGGGCATGGCCACCCGGGTGCCAA 662
QY 218 nGlnAsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLe 238
DB 663 CCAGAACCGCTTCTGCGCGACTGGAGACCCAGCGCCGCTGTGCTGTCCAGGCGCCCTGCC 722
QY 238 uAlaAlaArgSerHisSerSerTrpAsnSerAlaPhe 250
DB 723 ACCCTCAGGGGTGCGAGTCCACAACAGTGCCTTC 759
RESULT 7
US-08-167-628-1
: Sequence 1, Application US/08167628
: Patent No. 5408040
: GENERAL INFORMATION:
: APPLICANT: Grotendorst, Gary R.
: APPLICANT: Bradham Jr., Douglas M.,
: TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Spensley Horn Jubas & Lubitz
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/167,628
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/752,427
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Wetherell, Jr. Ph.D., John W.
: REGISTRATION NUMBER: 31,678
: REFERENCE/DOCKET NUMBER: PD-1294
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-455-5100
: TELEFAX: 619-455-5110
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2075 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: DB60R32
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 130..1177
: US-08-167-628-1
Alignment Scores: 7.41e-36 Length: 2075
Pred. No.: 7.41e-36

Score:	564.00	Matches:	112
Percent Similarity:	58.58%	Conservative:	28
Best Local Similarity:	46.86%	Mismatches:	81
Query Match:	39.17%	Indels:	18
DB:	1	Gaps:	6

US-10-010-408-2 (1-250) x US-08-167-628-1 (1-2075)

QY	13	SerpheLeu	CysLeuLeuSerMet-----Val	CysAlaGlnLeu	CysArgThr	28	
		::: :::	:::				
Db	163	GCCTTCGTGTCCTCCTCGCCCTCTG	CAAGCCGGCCGCGCTGCGGCAGAACTG	CAAGCGGG		222	
QY	29	ProCysThr	CysProTrrpThrPro---Pro	GlnCysPro	GlnGlyValPro	LeuValleu	47
					:::		
Db	223	CCGTGCGCGTGC	CCGAGACGACCGCGCGCGCTG	CCCGGGCGGCGTGA	CGCTGAGCTTCGTGCTG		282
QY	48	AspGlyCys	GlyCysCysLysValCysAlaArg	ArgLeuGlnGly	LeuSerCysAsp	HisLeu	67
Db	283	GACGCGCTGCGGCTGCTGCCGCGT	CGGCCCAAGCAGCTGGCGGAGCTGTG	CAACCGACCGC			342
QY	68	HisValCys	AspProSerGlnGlyLeuValCys	GlnProGlyAlaGlyPro	GlyGlyHis		87
Db	343	GACCCCTGCGACCCGCA	CAAGGGCCTCTTCTGTGACTTCGGCTCCCGGCCCAACCGCAAG				402
QY	88	GlyAlaVal	CysLeuLeuAspGlnAspAspGlySer	CysGlnValAsnGln	ArgArgTyr		107
Db	403	ATCGGCGTGTGACCGCC--AAAGAT	GTGTCTCCCTGCATCTTCGGTGTACGGTGTAC				459
QY	108	LeuAspGly	GlnThrPhelLysProAsnCysArgVal	LeuGlyCysArgCysAsp	AspGlyGly		127
Db	460	CGCAGCGGAGAGTCTCTTCCAGAG	CACTGCAAGTACAGTGCACGTGCCTG	CAAGCGGGCG			519
QY	128	PhethrCys	LeuProLeuCysSerGlnAspValArg	LeuProSerTrrpAspCysPro	Arg		147
Db	520	GTGGGCTGCATGCCCTGTGCA	GCATGGACGTTGCTGTGCCACAGCCCTGACTGCCCTTC				579
QY	148	ProLysArg	IleGlnValProGlyLysCysCysPro	GlnTrrpValCysAsp	GlnGlyVal		167
Db	580	CCGAGGAGGGTCAAGCTGCCCGGGA	ATGCTCTCGAGGAGTGGGTGTGACGAG-----				633
QY	168	ThrProAla	IleGlnArgSerThrAlaGlnGlyHis	GlnLeuSerAlaLeu-----			184
Db	634	---CCC	AAGACCAA-----ACCGTGTGGGCTGCCGCTGCGGCTTACCGACTGGAA				684
QY	185	-----Val	ThrProAlaSerAlaAspAlaPro	GlyProAsnTrrpSerThrAla			200
Db	685	GACACGTTTGGCCAGACCA	CTATGATTAGAGCCAACCTGCTGTCCAGACCACAGAG				744
QY	201	TrrpGlyPro	CysSerThrThrCysGlyLeuGlyIle	AlaThrArgValSerAsn	GlnAsn		220
Db	745	TGGAGCGGCTGTTC	CAAGACCTGTGGGATGGGCATCTCCACCCGGGTTACCA	TGACAAC			804
QY	221	ArgPheCys	GlnLeuGlnIleGlnArgArgLeu	CysLeuProArgPro	CysLeuAla		239
Db	805	GCCTCCTGCAGGCTAGAGAG	CAAGCCGCGCTGTGCATGTGTCAAGGCGCTTGCGCAAGCT				861

RESULT 8
US-08-386-680-1

```

: GENERAL INFORMATION:
: APPLICANT: Grotendorst, Gary R.
: APPLICANT: Bradham Jr., Douglas M.,
: TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Spensley Horn Jubas & Lubitz
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
:

```

```

1 MEDIUM TYPE: Floppy disk
2 COMPUTER: IBM PC compatible
3 OPERATING SYSTEM: PC-DOS/MS-DOS
4 SOFTWARE: Patentin Release #1.0, Ver
5 CURRENT APPLICATION DATA:
6 APPLICATION NUMBER: US/08/386,680
7 FILING DATE: 10-FEB-1995
8 CLASSIFICATION: 435
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: US/08/167,628
11 FILING DATE:
12 APPLICATION NUMBER: US/07/752,427
13 FILING DATE:
14 ATTORNEY/AGENT INFORMATION:
15 NAME: Wetherell, Jr. Ph.D., John W
16 REGISTRATION NUMBER: 31,678
17 REFERENCE/DOCKET NUMBER: PD-1294
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE: 619-455-5100
20 TELEFAX: 619-455-5110
21 INFORMATION FOR SEQ ID NO: 1:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 2075 base pairs
24 TYPE: nucleic acid
25 STRANDEDNESS: single
26 TOPOLOGY: linear
27 MOLECULE TYPE: CDNA
28 IMMEDIATE SOURCE:
29 CLONE: DB60R32
30 FEATURE:
31 NAME/KEY: CDS
32 LOCATION: 130..1177
33 US-08-386-680-1

```

Alignment Scores:	
Pred. No.:	7.41e-36
Score:	564.00
Percent Similarity:	58.58%
Best Local Similarity:	46.86%
Query Match:	39.17%
DB:	1
Length:	2075
Matches:	112
Conservative:	28
Mismatches:	81
Indels:	18
Gaps:	6

US-10-010-408-2 (1-250) x US-08-386-680-1 (1-2075)

QY	13.	SerPheLeuCysLeuLeuSerMet-----ValCysAlaGlnLeuCysArgThr	28
		::: : ::: :	
Db	163	GCCTTCGTGTCTCTCTCGCCCTTCGACCGCGCGCCGTCCGACGAACATGACGCGG	222
QY	29	ProCysThrCysProTrpThrPro---ProGlnCysProGlnGlyValProLeuValLeu	47
		:::	
Db	223	CCGTGCGGTGCCCGGACGAGCGCGCGCGCTGCCCCGCGGCGGTGAGCCTCGTGTG	282
QY	48	AspGlyCysGlyCysCysLysValCysAlaArgArgLeuGlyIleSerCysAspHisLeu	67
		::: ::: :	
Db	283	GACGGCTGGCGCTGCTGCCGCTGCGCCACAGCAGCTGGGCGAGCTGTGCACCGAGCGC	342
QY	68	HisValCysAspProSerGlnGlyValCysGlnProGlyAlaGlyProGlyGlyHis	87
		::: :::	
Db	343	GACCCCTGCGACCCCGCACAGGGCCCTCTTCTGTGACTTCGGCTCCCGGCCAACCGGCAAG	402
QY	88	GlyAlaValCysLeuLeuAspGlnAspAspGlySerCysGluValAsnGlyArgArgTyr	107
		:::	
Db	403	ATCGGCGTGTGCACCGCC--AAAGATGTGTGCTCCCTGCATCTTCGGTGTACGGTGTAC	459
QY	108	LeuAspGlyGluThrPheLysProAsnCysArgValLeuCysArgCysAspAspGlyGly	127
		::: : :: :::	
Db	460	CGCAGCGGAGAGTCTTCACAGACGACTGCAAGTACACAGTGCACGTGCCTGGACGCGGCG	519
QY	128	PheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrpAspCysProArg	147
		:	
Db	520	GTTGGCTGCATGCCCTGTGTGCACATGAGACGTTTCGTTGCCACGCCCTGACTGCCCTTC	579
QY	148	ProLysArgIleGlnValProGlyLysCysCysProGluTrpValCysAspGlnGlyVal	167

Db 580 CCGAGGAGGTCACAGCTGCCCCGGGAATGCTGCCGAGAGTGGGTGTGTGACGAG----- 633
QY 168 ThrProIaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu----- 184
Db 634 ---CCCAAGACCAA-----ACCGTGGTTGGGCGCTGCCCTCGCGGCTTACCGACTGGAA 684
QY 185 -----ValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 685 GACACGTTTGGCCCCAGACCCAACTATGATTAGACCAACTGCTGTGTCACAGACCACAGAG 744
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 745 TGGAGCGCTGTTCACAGACCTGTGGATGGGCATCTCCACCCGGGTTACCAATGACAAC 804
QY 221 ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
Db 805 GCCTCGTCAGGCTAGAGAAGCAGAGCGCCCTGTGCATGTGTCAGGCGCTTGCGAAGCT 861

RESULT 9

US-08-459-717-1
; Sequence 1, Application US/08459717
; Patent No. 5770209
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,717
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,427
; FILING DATE: 30-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
; US-08-459-717-1

Alignment Scores:
Pred. No.: 7 41e-36 Length: 2075
Score: 564.00 Matches: 112
Percent Similarity: 58.58% Conservative: 28
Best Local Similarity: 46.86% Mismatches: 81

Query Match: 39.17% Indels: 18
DB: 1 Gaps: 6
US-10-010-408-2 (1-250) x US-08-459-717-1 (1-2075)

QY 13 SerPheLeuCysLeuLeuSerMet-----ValCysAlaGlnLeuCysArgThr 28
Db 163 GCCTTCGTGTCCTCTCCCTGCGCTTCGACGCCGGCGCGCTGCGGCGGCGGCAACTGCAGCGGG 222
QY 29 ProCysThrCysProTrpThrPro--ProGlnCysProGlnGlyValProLeuValLeu 47
Db 223 CCGTGCCGGTGCCCGGACGAGCGCGGCGCGCTGCGCGGCGGCGGTGAGCCTCGTGTG 282
QY 48 AspGlyCysGlyCysCysLysValCysAlaArgArgLeuGlyGluSerCysAspHisLeu 67
Db 283 GACGGCTGGGCTGTGCTGCCGCTTCGCGCAAGCAGCTGGGCGAGCTGTGCACCGAGCGC 342
QY 68 HisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHis 87
Db 343 GACCCCTGGCAGCCCGCACAGAGGCGCTTCTGTGACTTCGGGCTCCCGGCCAACCGCAAG 402
QY 88 GlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluValAsnGlyArgArgTyr 107
Db 403 ATCGGCGGTGTGCACCGGCC---AAAGATGTGTCTCCCTGCATCTTCGGTGTACGGTGTAC 459
QY 108 LeuAspGlyGluThrPheLysProAsnCysArgValLeuCysArgCysAspAspGlyGly 127
Db 460 CGCAGCGGAGAGTCTCTCCAGAGCAGCTGCAAGTACAGTGCACGTGCTGACGGGGCGG 519
QY 128 PheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrpAspCysProArg 147
Db 520 GTGGGTGCATGCCCCCTGTGCAGCATGACGTTCGTGCCCCAGCGCTGACTGCCCTTTC 579
QY 148 ProlysArgIleGlnValProGlyLysCysCysProGluTrpValCysAspGlnGlyVal 167
Db 580 CCGAGGAGGTCACAGCTGCCCGGGAATGCTGCCAGAGAGTGGGTGTGTGACGAG----- 633
QY 168 ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu----- 184
Db 634 ---CCCAAGACCAA-----ACCGTGGTTGGGCGCTGCCCTCGCGGCTTACCGACTGGAA 684
QY 185 -----ValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 685 GACACGTTTGGCCCCAGACCCAACTATGATTAGACCAACTGCTGTGTCACAGACCACAGAG 744
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 745 TGGAGCGCTGTTCACAGACCTGTGGATGGGCATCTCCACCCGGGTTACCAATGACAAC 804
QY 221 ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
Db 805 GCCTCGTCAGGCTAGAGAAGCAGAGCGCCCTGTGCATGTGTCAGGCGCTTGCGAAGCT 861

RESULT 10

US-08-712-302-1
; Sequence 1, Application US/08712302
; Patent No. 5783187
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

QY 168 ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu----- 184
 DB 634 ---CCCAAGACCAA-----ACCGTGGTGGCCCTGCCTCGCGGCTTACCGACTGAA 684
 QY 185 -----ValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
 DB 685 GACACGTTTGGCCCAAGACCACTATGATTAGAGCAACTGCCGTGGTCCAGACACAGAG 744
 QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
 DB 745 TGGAGCGGCTGTCCAGACCTGTGGATGGGCATCTCCACCCGGGTTACCAATGACAAC 804
 QY 221 ArgPheCysGlnLeuGluIleGlnArgLeuCysLeuProArgProCysLeuAla 239
 DB 805 GCCTCTGACGCTAGAGAAGACAGCCCTGTGCATGTGTACAGGCTTGCGAAGCT 861

RESULT 13 US-09-080-715-1

; Sequence 1, Application US/09080715
 ; Patent No. 6190884
 ; GENERAL INFORMATION:
 ; APPLICANT: Grotendorst, Gary R.
 ; APPLICANT: Bradham Jr., Douglas M.
 ; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Spensley Horn Jubas & Lubitz
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/080,715
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/167,628
 ; FILING DATE:
 ; APPLICATION NUMBER: US/07/752,427
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wetherell, Jr. Ph.D., John W.
 ; REGISTRATION NUMBER: 31,678
 ; REFERENCE/DOCKET NUMBER: PD-1294
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-455-5100
 ; TELEFAX: 619-455-5110
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2075 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; IMMEDIATE SOURCE:
 ; CLONE: DB60R32
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 130..1177
 ; US-09-080-715-1

Alignment Scores:
 Pred. No.: 7.41e-36 Length: 2075
 Score: 564.00 Matches: 112
 Percent Similarity: 58.58% Conservative: 28
 Best Local Similarity: 46.86% Mismatches: 81

Query Match: 39.17% Indels: 18
 DB: 4 Gaps: 6

US-10-010-408-2 (1-250) x US-09-080-715-1 (1-2075)

QY 13 SerPheLeuCysLeuLeuSerMet-----ValCysAlaGlnLeuCysArgThr 28
 DB 163 GCCTTCGTGTCCTCTCCCTCGCCCTGTGACGCCGCCGCCGCGCTGCGCCAGAACTGCAGCGG 222
 QY 29 ProCysThrCysProTrpThrPro---ProGlnCysProGlnGlyValProLeuValLeu 47
 DB 223 CCGTCCCGGTGCCCCGAGCAGCGCGCGCGCGCTGCCCCGGCGGCTGAGCCTCGTGTG 282
 QY 48 AspGlyCysGlyCysCysLysValCysAlaArgArgLeuGlyGlnSerCysAspHisLeu 67
 DB 283 GACGGCTGGGCTGCTGCGCGCTGTGCGCAAGCAGCTGGGCGAGCTGTGCACCGAGCGC 342
 QY 68 HisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHis 87
 DB 343 GACCCCTGCGACCCGACACAGGCGCTCTGTGACTTCCGCTCCCGGCCCAACCGCAAG 402
 QY 88 GlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluValAsnGlyArgArgTyr 107
 DB 403 ATCGCGGTGTCACCGCC---AAAGATGTGTCTCCCTGCATCTTCGTGTGACGTGTAC 459
 QY 108 LeuAspGlyGluThrPheLysProAsnGlyValLeuCysArgCysAspAspGlyGly 127
 DB 460 CGCAGCGGAGATCCTCCAGAGCAGCTGCAAGTACAGTGCACAGTGCCTGACGCGGCG 519
 QY 128 PheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrpAspCysProArg 147
 DB 520 GTGGCTGATGCCCTGTGACAGCATGCATGTCTGCGCCAGCCCTGACTGCCCTTC 579
 QY 148 ProLysArgIleGlnValProGlyLysCysCysProGluTrpValCysAspGlnGlyVal 167
 DB 580 CCGAGGAGGTCAGCTGCCCGGGAATGCTGCGAGAGTGGTGTGACGAG----- 633
 QY 168 ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu----- 184
 DB 634 ---CCCAAGACCAA-----ACCGTGGTGGCCCTGCCTCGCGGCTTACCGACTGAA 684
 QY 185 -----ValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
 DB 685 GACACGTTTGGCCCAAGACCACTATGATTAGAGCAACTGCGTGGTCCAGACACAGAG 744
 QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
 DB 745 TGGAGCGGCTGTCCAGACCTGTGGATGGGCATCTCCACCCGGGTTACCAATGACAAC 804
 QY 221 ArgPheCysGlnLeuGluIleGlnArgLeuCysLeuProArgProCysLeuAla 239
 DB 805 GCCTCTGACGCTAGAGAAGACAGCCCTGTGCATGTGTACAGGCTTGCGAAGCT 861

RESULT 14 US-09-142-569-7

; Sequence 7, Application US/09142569
 ; Patent No. 6413735
 ; GENERAL INFORMATION:
 ; APPLICANT: Lau, Lester F.
 ; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30


```
Db      283  |||||||:|||||:|||||:|||||:|||||  ||
      GACGGCTGGCTGCTGCCGCGTGGCCAAAGCAGCTGGGCGAGCTGTGCACCGAGCGC 342
QY      68  HisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHis 87
      |||||||:|||||:|||||:|||||:|||||  ||
Db      343  GACCCCTGGGACCCCGACAAAGGGCTCTTGTGACTTCGGCTCCCGGCCAAACCGAAG 402
QY      88  GlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluValAsnGlyArgArgTyr 107
      |||||||:|||||:|||||:|||||:|||||  ||
Db      403  ATCGCGGTGGCACACCGCC--AAAGATGTGCTCCCTGCATCTTCGGTGTGTACGGTGTAC 459
QY      108  LeuAspGlyGluThrPheLysProAsnCysArgValLeuGlyArgCysAspAspGlyGly 127
      |||||||:|||||:|||||:|||||:|||||  ||
Db      460  CGCAGCGGAGAGTCCCTTCCAGAGCAGCTGCAAGTACCAGTGCACGCTGCTGACGGGGCG 519
QY      128  PheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrpAspCysProArg 147
      |||||||:|||||:|||||:|||||:|||||  ||
Db      520  GTGGGTGCATGCCCTGTGCAGATGGACGTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 579
QY      148  ProLysArgIleGlnValProGlyLysCysCysProGluTrpValCysAspGlnGlyVal 167
      |||||||:|||||:|||||:|||||:|||||  ||
Db      580  CCGAGGAGGCTCAAGCTGCCCGGAAATGCTGCGAGAGTGGGTGTGTGACGAG----- 633
QY      168  ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu----- 184
      |||||||:|||||:|||||:|||||:|||||  ||
Db      634  ---CCCAAGACCAA-----ACCGTGTGTGGGCTGCCCTGCCGGCTTACCGACTGGA 684
QY      185  -----ValThrProAlaSerAlaAspAlaProGlyProAsnTrpSerThrAla 200
      |||||||:|||||:|||||:|||||:|||||  ||
Db      685  GACACGTTGGCCCAAGACCAACTATGATTAGAGCCAACTGGCTGTCCAGACCAAGAG 744
QY      201  TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
      |||||||:|||||:|||||:|||||:|||||  ||
Db      745  TGGAGCGGCTGTTCAGAGACCTGTGGGATGGGCATCTCCACCCTGGTTTACCAATGACAAC 804
QY      221  ArgPheCysGlnLeuGluIleGlnArgLeuCysLeuProArgProCysLeuAla 239
      |||||||:|||||:|||||:|||||:|||||  ||
Db      805  GCGTCTGTCAGGCTAGAGAGCAGAGCGCTGTGCATGTGTACAGCGCTTGCGAAGCT 861
```

Search completed: August 1, 2003, 04:22:58
Job time : 59 secs

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 753 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..750
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-010-408-3

Alignment Scores:
Pred. No.: 4.26e-125 Length: 753
Score: 1440.00 Matches: 250
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-010-408-2 (1-250) x US-10-010-408-3 (1-753)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
DB 1 ATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCCACTTCGTTCCCTCTGCTCTCAATG 60
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
DB 61 GTGTGTGCCAGCTGTGTCCGACACCTGTGTAACCTGTCCTGACACCAACCCAGTGCCCA 120
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60
DB 121 CAGGGGTACCCCTGTGTGCTGATGCTGTGCTGTAAGTGTGTGCACGAGGCTG 180
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
DB 181 GGGGAGTCTCTCGACACCACTGCATGTCGCGACCCAGCCAGGGCCCTGTTGTCAGCCT 240
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
DB 241 GGGGCGAGCCCTGGCGCCATGGGGCTGTGTCTCTTGATGAGATGACGGTAGCCTGT 300
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
DB 301 GAGGTGAATGGCCGACGAGTACCTGATGAGAGAACCTTTAAACCAATTGACAGGTCCTG 360
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
DB 361 TGCCGCTGTGATGACGGTGGCTTCACTGCTGCGCTGTGACGTGAGATGTGCGGCTG 420
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
DB 421 CCCAGCTGGGACTGCCACGCCCCAAGAGAATACAGGTGCCAGGAAGTGTGCCCGGAG 480
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
DB 481 TGGGTATGTGACCAAGGAGTGACACCGCGCATCCAGCGGTCCACGGCGCAAGGACACCAA 540
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200

DB 541 CTTCTGCCCTGTGCACCTCCCTGCTGCTGATGCTCCTGTGTCCAAATTGGACACAGCC 600
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
DB 601 TGGGGCCCCCTGCTCAACCAACCACTGTGGCGCTGGGCATAGCCACCCGAGTGTCCAACGAGAC 660
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
DB 661 CGATTCTGCCAAGTGAAGATCCACAGCCCGCTGTGTCTGCCACAGACCTGCTGGCAGCC 720
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
DB 721 AGAGCCACAGCTCATGGAACAGTGTCTTC 750

RESULT 2

US-10-010-408-1
Sequence 1, Application US/10010408
Publication No. US20020165185A1
GENERAL INFORMATION:
APPLICANT: John J. Castellet, Jr.
TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules and Uses Therefor

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1708 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 249..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-010-408-1

Alignment Scores:
Pred. No.: 1.05e-124 Length: 1708
Score: 1440.00 Matches: 250
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-010-408-2 (1-250) x US-10-010-408-1 (1-1708)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db 249 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCCTTCTCTCTCTCAATG 308
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
Db 309 GTGTGTGCCACGCTGTGCGGACACCCCTGTACTGTCTTGGACACACCCAGTGCCTCA 368
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeu 60
Db 369 CAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGTAAGTGTGTGCACGGAGGCTG 428
QY 61 GlyGlySerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 429 GGGAGTCTCTGGACCACTGCATGTCTGGACCCACGCAAGGCTGGTTGTCAAGCT 488
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 489 GGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 548
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db 549 GAGGTGAATGGCCGACGACTACCTGGATGGAGAGACCTTAAACCCCAATTGACAGGGTCTCG 608
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 609 TGCCGCTGTGATGACGGTGGCTTACCTGCCTGCGCTGTGACAGTGAAGATGGCGCTG 668
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 669 CCCAGCTGGACCTGCCACGCCCAAGAAATACAGGTGCCAGAAAGTGTGCCCGAG 728
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 729 TGGGTATGTGACACGAGGTGACACCGGCGATCCAGCGCTCCACGGCGCAGACACCAA 788
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 789 CTTTCTGCCCTGTGACTCCTGCTCTGTGATGCTCTCTTCCAATTGGAGACACAGCC 848
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 849 TGGGGCCCTGCTCAACCACTGTGGCTGGGCATAGCCACCCGAGTGTCCAACAGAAAC 908
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 909 CGATTCTGCCAACTGGAGATCCAAACGCCGCTGTGTCTGCCAGACCCCTGCTGGCAGCC 968
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 969 AGGAGCCACAGCTCATGSAACAGTGTTC 998

RESULT 3
US-10-010-408-12

; Sequence 12, Application US/10010408

; Publication No. US20020165185A1

; GENERAL INFORMATION:

; APPLICANT: John J. Castelliott, Jr.

; TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules

; and Uses Therefor

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; ;

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/010,408
; FILING DATE: 07-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/044,273
; FILING DATE: March 19, 1998
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MBI-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..681
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-010-408-12

Alignment Scores:
Pred. No.: 2,83e-114 Length: 681
Score: 1323.00 Matches: 227
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 91.88% Indels: 0
DB: 15 Gaps: 0

US-10-010-408-2 (1-250) x US-10-010-408-12 (1-681)
QY 24 GlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysProGlnGlyVal 43
Db 1 CAGCTGTGCCGACACCTGTACTGTCTTGGAGACACCAACCCAGTGCACAGGGGTA 60
QY 44 ProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeuGlyGlySer 63
Db 61 CCCGTGCTGTGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGAGGCTGGGGAGTCC 120
QY 64 CysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGly 83
Db 121 TGCAGCACCTGATGTCTGCGACCCACGAGGGCTGTGTGTGTACGCTGGGGCAGGC 180
QY 84 ProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluValAsn 103
Db 181 CTTGGCGGCATGGGGCTGTGTGTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT 240
QY 104 GlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeuCysArgCys 123
Db 241 GGCCGAGGTACTGGATGGAGAGACCTTAAACCCCAATTGCAGGGTCTGTGCCGCTGT 300
QY 124 AspAspGlyGlyPheThrCysLeuProLeuCysSerGlnGlyLysCysCysProGluTrpValCys 143
Db 301 GATGACGGTGGCTTACCTGCCCTGCCGCTGTGAGATGAGATGTGCGGCTGCCACAGCTGG 360
QY 144 AspCysProArgProLysArgIleGlnValProGlyLysCysCysProGluTrpValCys 163
Db 361 GACTGCCACGCCCCAAGAATATACAGGTGCCAGGAAGTGTGCCCCAGTGGGTATGT 420
QY 164 AspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAla 183
Db 421 GACGAGGAGTGAACCGCGGATCCAGCGCTCCACGGCGCAAGGACACCACTTCTGCC 480
QY 184 LeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAlaTrpGlyPro 203

```
Db      481 CTTGTCACCTCTGCTCTGATGCTCTTGTCCAAATGGAGCACAGCCCTGGGCCCC 540
QY      204 CysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsnArgPheCys 223
Db      541 TGGTCAACACACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAACCAGAACTTCTGC 600
QY      224 GlnLeuGlnIleGlnArgArgLeuCysLeuProArgProGlySerLeuAlaAlaArgSerHis 243
Db      601 CAACGTGAGATCCAAACGCGCCTGTGTCTGCCCCAGACCCCTGGCTGGCAGCCAGAGCCAC 660
QY      244 SerSerTrpAsnSerAlaPhe 250
Db      661 AGCTCATGGAACAGTGTCTTC 681
```

RESULT 4

```
US-10-112-267-17
; Sequence 17, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-112-267-17
```

Alignment Scores:

```
Pred. No.: 1.75e-112 Length: 1734
Score: 1308.50 Matches: 226
Percent Similarity: 93.63% Conservative: 9
Best Local Similarity: 90.04% Mismatches: 15
Query Match: 90.87% Indels: 1
DB: 15 Gaps: 1
```

US-10-010-408-2 (1-250) x US-10-112-267-17 (1-1734)

```
QY      1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet 20
Db      257 ATGAGGGCAACCACTGATCATCTTGTGGCATTTCTTCTCTGCAATCTCTCAATG 316
QY      21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
Db      317 GTGTATTCACAGCTGTGCCACAGACCCCTGTGCTTCCCTTGACACACACCCAGTGCCCA 376
QY      41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaAlaArgLeu 60
Db      377 CCGGGGGTACCCCTGTGCTGATGGCTGTGGCTGCTGTCGAGTGTGTGCACGGAAGGCTG 436
QY      61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
```

```
Db      437 GGGGAGTCCTGCGACCACTGCATGTCCTGCGACCCCAAGCCAGGGCCCTGGTTGTACAGCT 496
QY      81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlnAspArgPheCys 100
Db      497 GGGGACAGCCCGCAGTGGCCGTGGTGTCTGTGTGCTCTTCCAAAGAGATGACGGAGCTGT 556
QY      101 GluValAsnGlyArgArgTyrLeuAspGlyGlnThrPheLysProAsnCysArgValLeu 120
Db      557 GAGGTGAATGGCCCGCAGTACCTGTGATGGGAGACCTTTAAACCAATTCAGAGGTTTGTG 616
QY      121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLeu 140
Db      617 TGCCGCTGTGATGACGGTGGTTTCACCTGCTGCGCGCTGTGCAGTGAAGATGTGCGGCTG 676
QY      141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db      677 CCCAGCTGGGACTGCCCCACGCCCCAGAGAGATACAGGTGCCAGAGAGTGTGCCCGGAG 736
QY      161 TrpValCysAspGlnGlyVal---ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179
Db      737 TGGGTGTGTGACCAAGCAGTGTATGACGCCGCGCAATCCAGCCCTCTCAGCCCAAGACAC 796
QY      180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199
Db      797 CAACCTTCTGCGCTGTGACCTCTGATGATGCGCATGCGCATGCGCCCTGTCCAACTGGAGACA 856
QY      200 AlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln 219
Db      857 GCCTGGGGCCCTGTGCAACCAACCTGTGGGTGGGCATAGCCACCCGAGTATCCAAACCAG 916
QY      220 AsnArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
Db      917 AACCGATTCTGCCAAGTGAAGATCCAGCGTGGCTGTGTCTGTCCAGACCCCTGCGTGA 976
QY      240 AlaArgSerHisSerSerTrpAsnSerAlaPhe 250
Db      977 TCCAGAGACCAAGGCTCATGGAACAGTGCCTTC 1009
```

RESULT 5

```
US-10-112-267-18/c
; Sequence 18, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 18
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-112-267-18
```

Alignment Scores:

Pred. No.: 1.75e-112 Length: 1734
Score: 1308.50 Matches: 226
Percent Similarity: 93.63% Conservative: 9
Best Local Similarity: 90.04% Mismatches: 15
Query Match: 90.87% Indels: 1
DB: 15 Gaps: 1

US-10-010-408-2 (1-250) x US-10-112-267-18 (1-1734)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db 1478 ATGAGGGGCAACCCACTGATCCATCTTGCCATTTCCTTCCCTGCAATCTCTCAATG 1419
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
Db 1418 GTGTATTCACAGCTGTGCCAGCACCCCTGTGCTGTCTTGACACACCACCCAGTGCCCA 1359
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeu 60
Db 1358 CCGGGGTACCCCTGTGCTGTGATGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1299
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 1298 GGGAGTCCCTGGCAGCACCTGCATGTCTGCCAGCCCAAGCAGGGCCCTGGTTTGTCAAGCT 1239
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 1238 GGGGAGGCCCCAGTGGCCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1179
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db 1178 GAGGTGAATGGCCGACAGGTACCTGGATGGGGAGACCTTTAAACCCCAATTGCAGGGTTTGG 1119
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 1118 TGCCGCTGTGATGAGGTGTGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1059
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 1058 CCCAGCTGGAGTGGCCACGCCCCAGAGAAATACAGGTGCCAGGAGAGTGTGCCCCGAG 999
QY 161 TrpValCysAspGlnGlyVal--ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179
Db 998 TGGGTGTGTGACCAGGAGCATGTATGCAGCCGGAATCCAGCCCTCTCAGCCCCAAGGACAC 939
QY 180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199
Db 938 CAACCTTCTGCCCTTGTCTACTCTCTGATCTGTCCGATGGCCCCCTGTCCAACCTGGAGACA 879
QY 200 AlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln 219
Db 878 GCCTGGGGCCCCCTGCTCAACACACCTGTGGGTGGGCATAGCACCCGAGTATCCAAACAG 819
QY 220 AsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239
Db 818 AACCGATTCTGCCAACTGGAGATCCAGCGTGCCTGTGTCTGTCCAGACCCCTGCTGCA 759
QY 240 AlaArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 758 TCCAGGAGCCACGGCTCATGGAACAGTGCCTTC 726

RESULT 6

US-10-137-866-319
; Sequence 319, Application US/10137866
; Publication No. US20030129689A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C151
CURRENT FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03

Db 250 GGGGACGACCCGGTGGCCGGGGGGCCCTGTGCTCTTGGCAGAGAGACGACAGCAGCTGT 309
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db 310 GAGGTGAACGGCCCGCTGTATCGGAGAGGGAGACCTTCAGCCCGCAGCATCCGC 369
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 370 TGCCGCTGCGAGGACGGCGGCTTCACCTGCGCTGTGCGCAGAGATGTGCGGCTG 429
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 430 CCCAGCTGGGACTGCCCCACCCAGAGGGTGCAGAGTCCGTGGGCAAGTGTGCCCTGAG 489
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGGCGCCCAAGAGGGGAGCTGGGAGACCCCTTCAGCCCAAGAGACCCCGAG 549
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCCTTGTCTCTTCCCTGCCCTGTGTGTCCTGCCCGCAGATGAGCAGCGGCC 609
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGACCTGTGCTGCAGCACCTGTGGGCTGGGCATGGCACCCGGGTGTCCAACACAGAAC 669
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 670 CGCTTCTGCCGACTGCAGACCCAGCGCCCTGTGCCCTGTCCAGGGCCCTGCCACCCCTCC 729
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 730 AGGGGTGCGCAGTCCACAAACAGTGCCTTC 759

RESULT 7

US-10-146-726-319
; Sequence 319, Application US/10146726
; Publication No. US20030129690A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C308
; CURRENT APPLICATION NUMBER: US/10/146,726
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-726-319

Alignment Scores:
Pred. No.: 6.44e-90 Length: 1266
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17

Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 14 Gaps: 0
US-10-010-408-2 (1-250) x US-10-146-726-319 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20
Db 10 ATGAGAGGCACACCGAAGACCACCTCCTGCGCTTCCCTCCTCCTCCTCCTCAAG 69
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
Db 70 GTGCGTACCAGCTGTGCCCGACACCATGTACCTGCCCTGCCACACCTCCCGATGCCG 129
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60
Db 130 CTGGGAGTACCCCTGTGCTGTGATGGCTGTGGCTGCTGCGGGGTATGTGCACGGCGCTG 189
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 190 GGGGACCTGTGCGACCAACTCCACGTCTGCGAGCGCCAGCCAGGGCCCTGTGTGCCAGCCC 249
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlyAspAspGlySerCys 100
Db 250 GGGCAGAGACCCCGGTGGCGCGGGGCCCTGTGCTCCTGTGGCAGAGACGACAGCAGACTGT 309
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db 310 GAGGTGAACGGCCCGCTGTATCGGAGAGGGAGAGACCTTCAGCCCGCAGCATCCGC 369
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 370 TGCCGCTGCGAGGACGGCGGCTTCACCTGCGCTGCCGTGTGCAGCGAGGATGTGCCGCTG 429
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 430 CCCAGCTGGGACTGCCCCACCCAGAGGGTGCAGAGTCCGTGGCAGAGTGTGCCCTGAG 489
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGGCGCCCAAGAGGGGAGCTGGGAGACCCAGCCCTTCAGCCCAAGAGACCCCGAG 549
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCCTTGTCTCTTCCCTGCCCTGTGTGTCCTGCCCGCAGATGAGCAGCGGCC 609
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGACCTGTGCTGCAGCACCTGTGGGCTGGGCATGGCACCCGGGTGTCCAACACAGAAC 669
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 670 CGCTTCTGCCGACTGCAGACCCAGCGCCCTGTGCTGTCTTCAGGGCCCTGCCACCCCTCC 729
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 730 AGGGGTGCGCAGTCCACAAACAGTGCCTTC 759

RESULT 8

US-10-146-727-319
; Sequence 319, Application US/10146727
; Publication No. US20030129691A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT:


```

; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C312
; CURRENT APPLICATION NUMBER: US/10/146,727
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-727-319

```

Alignment Scores:	
Pred. No.:	6.44e-90
Score:	1064.00
Percent Similarity:	80.40%
Best Local Similarity:	73.60%
Query Match:	73.89%
DB:	14
Length:	1266
Matches:	184
Conservative:	17
Mismatches:	49
Indels:	0
Gaps:	0

US-10-010-408-2 (1-250) x US-10-146-727-319 (1-1266)

QY	1	Met	arg	Gly	Ser	Pro	Leu	Ile	His	Leu	Leu	Ala	Thr	Ser	Ph	Leu	Cys	Leu	Ser	Met	20		
Db	10	ATG	AGAGG	CAC	ACC	CGA	AGACCA	CTC	TGG	CCCTT	CTG	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	69		
QY	21	Val	Cys	Ala	Gln	Leu	Cys	Arg	Thr	Pro	Cys	Thr	Cys	Pro	Trp	Thr	Pro	Pro	Gln	Cys	Pro	40	
Db	70	GTC	GTA	CC	CAG	CTG	TG	CCCG	AC	CA	TG	TAC	CTG	CCCG	TG	GC	CA	CC	TCC	CCG	ATG	129	
QY	41	Gln	Gly	Val	Pro	Leu	Val	Leu	Asp	Gly	Cys	Gly	Cys	Ser	Val	Cys	Ala	Arg	Arg	Leu	60		
Db	130	CTG	GAG	AGT	AT	CC	CTG	TG	CTG	TGA	TG	CGT	G	TG	CGT	G	TG	CGG	G	TAT	GC	189	
QY	61	Gly	Glu	Ser	Cys	Asp	His	Leu	His	Val	Cys	Asp	Pro	Ser	Gln	Gly	Leu	Val	Cys	Gln	Pro	80	
Db	190	GGG	GAG	CC	CTG	CGA	CTC	CA	CTG	CTG	CGA	CGC	AG	CGA	G	GGC	CTG	TG	CC	CA	GGC	249	
QY	81	Gly	Ala	Gly	Pro	Gly	Gln	His	Gly	Ala	Val	Cys	Leu	Leu	Asp	Glu	Asp	Asp	Gly	Ser	Cys	100	
Db	250	G	G	G	C	A	G	A	C	C	G	G	T	G	G	C	C	C	T	G	C	T	309
QY	101	Glu	Val	Asn	Gly	Arg	Arg	Tyr	Leu	Asp	Gly	Glu	Thr	Ph	ely	S	Pro	Asn	Cys	Arg	Val	Leu	120
Db	310	GAG	GTG	AAC	G	G	C	CG	CTG	TAT	G	G	G	GA	A	G	G	G	AG	AC	CTT	GC	369
QY	121	Cys	Arg	Cys	Asp	Asp	Gly	Gly	Ph	eth	Thr	Cys	Leu	Pro	Leu	Cys	Ser	Glu	Asp	Val	Arg	Leu	140
Db	370	TG	CCG	CTG	C	G	A	G	A	C	G	G	CTT	C	A	CTG	C	G	TG	CCG	CTG	TG	429
QY	141	Pro	Ser	Trp	Asp	Cys	Pro	Arg	Pro	Lys	Arg	Ile	Gln	Val	Pro	Gly	Lys	Cys	Cys	Pro	Glu	160	
Db	430	CCC	AGC	TG	G	ACT	G	CCCC	CCCC	AC	CC	CA	GGA	G	GGT	C	AG	AGT	CTG	G	G	CA	489
QY	161	Trp	Val	Cys	Asp	Gln	Gly	Val	Thr	Pro	Ala	Ile	Gln	Arg	Ser	Thr	Arg	Gln	Gly	His	Gln	180	
Db	490	TGG	GTG	TG	CGG	CA	AG	G	GG	G	ACT	TG	G	G	GA	CC	CA	CC	CTT	CC	AG	CC	549
QY																							

QY	221	ArphneCysGlnLeuCluIleGlnArgArgLeuCysLeuProArgProcCysLeuAlaAla	240
		:::	:::
Db	670	CgCTTCTGCCGACTGSAGAACCCAGCGCGCCCTGTGCCTTCACGAGCCCTGCCACCCCTCC	729
QY	241	ArgSerHisSerSerTrpAsnSerAlaphe	250
Db	730	AGGGGTGCGACAGTCCACAACAGTGCCCTTC	759

RESULT 9
US-10-146-788-319
; Sequence 319, Application US/10146788
; Publication No. US20030129692A1
; GENERAL INFORMATION:

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC322
; CURRENT APPLICATION NUMBER: US/10/146,788
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-788-319

```

Alignment Scores:	6.44e-90	Length:	1266
Pred. No.:	1064.00	Matches:	184
Score:	80.40%	Conservative:	17
Percent Similarity:	73.60%	Mismatches:	49
Best Local Similarity:	73.89%	Indels:	0
Query Match:	14	Gaps:	0
DB:			

US-10-010-408-2 (1-250) x US-10-146-788-319 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
 |||||:::||||| ||||| ||| |||||
 Db 10 ATGAGAGGCACACCGAAGACCACCTCTGCGCTTCTCCCTCTCTGCTCTCTCAAG 69
 QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
 ||| ||||| ||||| ||||| ||||| |||||:::|||||
 Db 70 GTGCGTACCCACAGCTGTGCCCGACACACATGTACTGCCCCCTGGCCACCTCCCGATGCCG 129
 QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
 ||||| ||||| ||||| ||||| |||||:::||||| |||||
 Db 130 CTGGGAGTACCCCTGGTGTGCTGATGGCTGTGGCTGCTGCCGGGTATGTGCACGGCGGCTG 189
 QY 61 GlyGlnSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 190 GGGGAGCCCTGGCGAACCAACTCCACGCTGTGCGAAGCCAGCCAGGGGCTGTGTCACAGCCC 249
 QY 81 GlyAlaGlyProGlyGlyGlnHisGlyAlaValCysLeuLeuAspGlnAspAspGlySerCys 100
 ||||| ||||| |||||:::||||| ||||| |||||
 Db 250 GGGGAGGACCCGGGTGGCCGGGGGGGCTGTGCTCTTGGCAGAGGAGACAGCAGCTGT 309

QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPhelysProAsnGlySerGlyValLeu 120
Db 310 GAGTGAACGGCGCGCTGTATCGGAAGGGAGACCTTCCAGCCCCACTGACATCCGC 369
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlyAspValArgLeu 140
Db 370 TGGCGCTGCGAGAGCGCGGCTTACCTGCGCTGCGCTGTGCAGCGAGCATGTGCGGCTG 429
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 430 CCCAGCTGGAGTGCCTCCACCCAGAGGGTGAAGTCTGTGGGCAAGTGCCTGAG 489
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGCGCCCAAGAGGGGAGTGGGACCCACCTTCCAGCCCAAGAGACCCAG 549
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCCTGTCTCTTCCCTGCCCGCTGTGCTGCCCTGAGAGAGACAGGCC 609
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGACCCGTGCTGCAGACCACTGTGGGCTGGGCATGGCCACCCGGGTGCCAACCAGAAC 669
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 670 CGCTTCTGCGACTGAGAGACCAGCGCGCTGTGCTGCCAGGGCCCTGCCACCCCTCC 729
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 730 AGGGGTGCGAGTCCACAACAGTGCCTTC 759

RESULT 10
US-10-152-380-319
; Sequence 319, Application US/10152380
; Publication No. US20030129694A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C397
; CURRENT APPLICATION NUMBER: US/10/152,380
; CURRENT FILING DATE: 2002-05-21
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-152-380-319

Alignment Scores:
Pred. No.: 6.44e-90
Score: 1064.00
Percent Similarity: 80.40%
Best Local Similarity: 73.60%

Query Match: 73.89% Indels: 0
DB: 14 Gaps: 0
US-10-010-408-2 (1-250) x US-10-152-380-319 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db 10 ATGAGAGGCACACCGAAGACCACCTCTGCGCTTCTCCCTCTGCTCTCTCAAG 69
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
Db 70 GTGCGTACCCTGCTGCGCCGACACCATGATCTGCCCTGCGCACCTCCCGATGCCCG 129
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
Db 130 CTGGAGTACCCCTGTGCTGTGATGGCTGTGCTGTGCCGGGTATGTGCACGGCGCTG 189
QY 61 GlyGlySerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 190 GGGAGCCCTGGACCACTCCACGTCTGCGAGCCAGCCAGGGCGCTGTGCCAGCCC 249
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlyAspAspGlySerCys 100
Db 250 GGGGACAGACCCGGTGTGGCCGGGGGGCTGTGCTCTTGTGGCAGAGAGACAGACAGCTGT 309
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPhelysProAsnGlySerGlyValLeu 120
Db 310 GAGTGAACGGCGCGCTGTATCGGGAAGGAGACCTTCCAGCCCACTGACATCCGC 369
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlyAspValArgLeu 140
Db 370 TGGCGTGCAGAGACGGCGCTTACCTGCTGCGCTGTGCAGCGAGATGTGCGGCTG 429
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 430 CCCAGCTGGAGTGCCTCCACCCAGAGGGTGAAGTCTGTGCCAGCGAGATGTGCGGCTG 489
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGCGCCCAAGAGGGGAGTGGGACCCACCCCTTCCAGCCCAAGAGACCCAG 549
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCCTGTCTCTTCCCTGCCCGCTGTGCTGCCCTGAGAGAGACAGGCC 609
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGACCCGTGCTGCAGACCACTGTGGGCTGGGCATGGCCACCCGGGTGCCAACCAGAAC 669
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 670 CGCTTCTGCGACTGAGAGACCAGCGCGCTGTGCTGCCAGGGCCCTGCCACCCCTCC 729
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 730 AGGGGTGCGAGTCCACAACAGTGCCTTC 759

RESULT 11
US-10-153-934-319
; Sequence 319, Application US/10153934
; Publication No. US20030129695A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven

APPLICANT: Smith,Victoria
APPLICANT: Stewart,Timothy A.
APPLICANT: Tumas,Daniel
APPLICANT: Watanabe,Colin K
APPLICANT: Wood,William
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C412
CURRENT APPLICATION NUMBER: US/10/153,934
CURRENT FILING DATE: 2002-05-22
Prior Application removed - See file Wrapper or Pajm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 319
LENGTH: 1266
TYPE: DNA
ORGANISM: Homo Sapien
US-10-153-934-319

Alignment Scores:
Pred. No.: 6 44e-90 Length: 1266
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 14 Gaps: 0

US-10-010-408-2 (1-250) x US-10-153-934-319 (1-1266)

OY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
DB 10 ATGAGAGGCAACACCGAAGACCACCTGCGCTTCTGCTCCCTGCTCCCTCAAG 69
OY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrThrProProGlnCysPro 40
DB 70 GTCCGTACCCAGCTGTGCCCGACACCATGTACTGCCCTGGCCACCTCCCGATGCCG 129
OY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysValCysAlaArgLeu 60
DB 130 CTGGAGTACCTGTGTGTGATGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189
OY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
DB 190 GGGAGCCCTGCGACCAACTCCAGCTGTGCGACGCCAGCGGCGCTGTGCCAGCCC 249
OY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlnAspAspGlySerCys 100
DB 250 GGGCAGAGCCGGGTGGCGGGGGCTGCTCTTGGCAGAGGACGACAGCAGCTGT 309
OY 101 GluValAsnGlyArgArgTyrLeuAspGlyGlyLysThrPheLysProAsnCysArgValLeu 120
DB 310 GAGGTGAACGGCCGCTGTATCGGAGAGGGGAGACCTTGCAGCCCCACTGCAGCATCCGC 369
OY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLeu 140
DB 370 TGCCGCTGCGAGAGCGGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429
OY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGln 160
DB 430 CCCAGCTGGGACTGCCCCACCCAGAGGGGTGAGGTGCTGGCAAGTGTGCCCTGAG 489
OY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
DB 490 TGGGTGTGGGCCCAAGAGGGGAGCTGGGACCCAGCCGCTTCCAGCCCAAGAGACCCAG 549
OY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
DB 550 TTTTCTGGCCTTGTCTCTTCCCTGCCCTGTGTCCCTGCTGCCAGATGAGACAGGCC 609
OY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
DB 610 TGGGACCTGCTGACCAACCTGTGGGTGGGATGGCGACCCGGGTGTCCAACAGAAC 669

OY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
DB 670 CGCTTCTGCCGACTGGAGACCAGCGCGCTGTGCTGTCCAGGCGCTGCCACCTCC 729
OY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
DB 730 AGGGGTCCGACTCCACAACAGAGTGCCTTC 759
RESULT 12
US-10-028-072-319
Sequence 319, Application US/10028072
Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327

; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24

; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Alignment Scores:

Pred. No.:	6.44e-90	Length:	1266
Score:	1064.00	Matches:	184
Percent Similarity:	80.40%	Conservative:	17
Best Local Similarity:	73.60%	Mismatches:	49
Query Match:	73.89%	Indels:	0
DB:	15	Gaps:	0

US-10-010-408-2 (1-250) x US-10-028-072-319 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet 20
Db 10 ATGAGAGGACACACCAGAACACCACCTCCTGGCCCTCTCCCTCCTCTGCTCTCAAG 69

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
Db 70 GTGCGTACCCAGCTGTGCCCCGACACCATGTACTGCCCCCTGGCCACCTCCCGATGCCCG 129

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgLeu 60
Db 130 CTGGAGTACCCCTGGTGTGCTGGATGGCTGTGGCTGTGCCGGTATGTGCAGCGCGCTG 189

QY 61 GlyGlySerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 190 GGGAGACCCCTGGACCAACTCCACGTCTGGAGCCAGCCAGGCGCTGGTCCAGCCGCC 249

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlyAspArgCys 100
Db 250 GGGGACAGACCCGGTGGCGCGGGGGCGCTGTGCTTGGCAGAGACGACAGCAGCTGT 309

QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGlyThrPheLysProAsnCysArgValLeu 120
Db 310 GAGGTGAACGGCCGCTGTATCGGGAGAGGGAGACCTTCCAGCCCACTGCACATCCGC 369

QY 121 CysArgCysAspArgGlyGlyPheThrCysLeuProLeuCysSerGlyAspValArgLeu 140
Db 370 TGCCGCTGCAGAGACGGCGGCTTCACTGCTGCGGCTGTGCAGCAGATGTGGGCTG 429

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 430 CCCAGCTGGGACTGCCCCCACCAGGAGGGTCTGAGGTCTGGGCAAGTGTGCCCTGAG 489

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGGCGCCCAAGAGGGGGGACTGGGAGCCAGCCCTTCCAGCCCAAGAGCCAG 549

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCCTGTCTCTCCCTGCCCCCTGGTGTCCCTGCCCCAGAAATGGAGCAGGCC 609

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGAGCCCTGCTGCACCACTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACAGAAC 669

QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 670 CGCTTCTGCCGACTGGAAGACCCAGCGCCGCTGTGCTGTCCAGGGCCCTGCCACCCCTCC 729

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 730 AGGGGTGCGAGTCCACAAAACAGTGCCTTC 759

RESULT 13
US-10-121-049-319
; Sequence 319, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-121-049-319

Alignment Scores:
Pred. No.: 6.44e-90 Length: 1266
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: Gaps: 0

US-10-010-408-2 (1-250) x US-10-121-049-319 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet 20
Db 10 ATGAGAGGACACACCAGAACACCACCTCCTGGCCCTCTCCCTCCTGCTCTCAAG 69

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40
Db 70 GTGCGTACCCAGCTGTGCCCCGACACCATGTACTGCCCCCTGGCCACCTCCCGATGCCCG 129

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgLeu 60
Db 130 CTGGAGTACCCCTGGTGTGCTGGATGGCTGTGCTGCTGCCGGTATGTGCAGCGCGCTG 189

QY 61 GlyGlySerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 190 GGGAGACCCCTGGACCAACTCCACGTCTGGAGCCAGCCAGGCGCTGGTCTGCAGGCC 249

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlyAspArgCys 100
Db 250 GGGGACAGACCCGGTGGCGCGGGGGCGCTGTGCTTGGCAGAGACGACAGCAGCTGT 309

QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGlyThrPheLysProAsnCysArgValLeu 120
Db 310 GAGGTGAACGGCCGCTGTATCGGGAGAGGGAGACCTTCCAGCCCACTGCACATCCGC 369

QY 121 CysArgCysAspArgGlyGlyPheThrCysLeuProLeuCysSerGlyAspValArgLeu 140
Db 370 TGCCGCTGCAGAGACGGCGGCTTCACTGCTGCGGCTGTGCAGAGAGATGTGGGCTG 429

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 430 CCCAGCTGGGACTGCCCCACCCAGAGGGTCTGAGTCTGGGCAAGTGTGCCCTGAG 489

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGGCGCCCAAGAGGGGAGCTGGGAGCCAGCCCTTCCAGCCCAAGAGCCAG 549

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCCTTGTCTCTCCCTGCCCCCTGGTGTCCCTGCCCCAGAAATGGAGCAGGCC 609

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGACCTGCTGCACCACTGTGGCTGGGCATGGCCACCCGGGTGTCCAACAGAAC 669

QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 670 CGCTTCTGCCGACTGGAAGACCCAGCGCCGCTGTGCTGTCCAGGCGCTGCCACCCCTCC 729

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 730 AGGGGTCCGAGTCCACAAACAGTGCCTTC 759

RESULT 14
US-10-123-904-319

; Sequence 319, Application US/10123904
; Publication No. US20030022328A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-904-319

Alignment Scores:
Pred. No.: 6.44e-90 Length: 1266
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 15 Gaps: 0

US-10-010-408-2 (1-250) x US-10-123-904-319 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db 10 ATGAGAGGACACACCGAAGACCCACCTCTGCGCTTCTCCCTCCTGCTCTCTCAAG 69
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
Db 70 GTGCGTACCACTGTGCGCCGACACCACTGACCTGCCCTGGCCACCTCCCGATGCCCG 129
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu 60
Db 130 CTGGAGTACCCCTGTGCTGTGATGGCTGTGCTGTGCCGGGTATGTGCACGGCGCTG 189
QY 61 GlyGlySerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 190 GGGAGCCCTGGACCACTCCACGTGTGCACGCCACGAGCCAGGGCCGTGTGCAAGCC 249
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlnAspAspGlySerCys 100
Db 250 GGGCAGGACCCGGTGGCGGGGGCCCTGTGCTCTTGGCAGAGGACGACGACGCTGT 309
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGlyLeuThrPheLysProAsnCysArgValLeu 120
Db 310 GAGGTGAACGGCCCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCACGATCCGC 369

QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 370 TGCCCGTCCGAGGACGGCGGCTTACCTGCGTGCCTGTGACAGGAGATGTGGCGCTG 429

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGln 160
Db 430 CCCAGCTGGAGCTGCCCCACCCACCCAGAGGTCGAGGTCTGGGCAAGTGTGCTGCTGAG 489

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGGCCCAAGAGAGGGGAGCTGGGACCCAGCCCTTTCAGCCCAAGACCCAG 549

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCCTGTCTCTTCCCTGCCCTGTGCTGCCCTGCCAGATGGACACGGCC 609

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGACCTGTGCTCGACCACTGTGGCTGGGCATGGCCACCCGGGTGTCCAACAGAAC 669

QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 670 CGCTTCTGCCGACTGGAAGACCCAGCGCGCTGTGCTGTCCAGGCCCTTGCCACCTCC 729

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 730 AGGGGTCCGAGTCCACAAACAGTGCCTTC 759

RESULT 15
US-10-140-470-319

; Sequence 319, Application US/10140470
; Publication No. US20030022331A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-470-319

Alignment Scores:
Pred. No.: 6.44e-90 Length: 1266
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 15 Gaps: 0

US-10-010-408-2 (1-250) x US-10-140-470-319 (1-1266)


```
OY      1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db      10 ATGAGAGGCACACCCAGAACCCACCCTCCTGGCTTCTCCCTCCCTGCTGCTCTCTCAAG 69
OY      21 ValCysAlaGlnLeuCysArgThrProCysThrCysProGlyThrProProGlnCysPro 40
Db      70 GTGCGTACCCAGCTGTGCCGACACACCATGACTGCCCCCTGGCCACCCTCCCGATGCCCG 129
OY      41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60
Db      130 CTGGGAGTACCCCTGGTGGCTGGATGGCTGTGGCTGCTGCCGGGGTATGTGCACGGCGGCTG 189
OY      61 GlyGlySerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db      190 GGGGAGCCCTGCGACCACTCCACGCTGTCGACGCCACGCCAGGGCCCTGGTCTGCCAGCCC 249
OY      81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db      250 GGGGAGACCCCGGTGGCCGGGGGGGGCCCTGTGCTGTCGACAGAGACGACAGCAGCTGT 309
OY      101 GluValAsnGlyArgArgTyrLeuAspGlyIuThrPheLysProAsnCysArgValLeu 120
Db      310 GAGGTGAACGGCCCGCTGTATCGGGAAGGGGAGACCTTCGAGCCCACTGCACGATCCCGC 369
OY      121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuGysSerGluAspValArgLeu 140
Db      370 TGCCGCTGCGAGAGCGGGGCTTCACCTGCGCTGCCGCTGTGCAGCGAGATGTCCGGCTG 429
OY      141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db      430 CCCAGCTGGGACTGCCCCCACCACCAGAGAGGCTCGAGTCTGGGCAAGTCTGCCCTGAG 489
OY      161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db      490 TGGGTGTCGGCCAGAGAGGGGACTGGGGAGCCACCCCTTCACAGCCCAAGGACCCAG 549
OY      181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProGysProAsnTrpSerThrAla 200
Db      550 TTTTCTGGCCTTGTCTCTTCCCTGCCCTGGTGTCCCTTGCCTCAGAAATGAGACAGGCC 609
OY      201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db      610 TGGGACCCCTGCTCGACCACTGTGGGTGCGCATGGCCACCCGGGTGTCCAACAGAAC 669
OY      221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db      670 CGCTTCTGCGGACTGGAGAGACCCAGCGCCGCTGTGCCCTGTCCAGGGCCCTGCCACCCCTCC 729
OY      241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db      730 AGGGGTGCGCAGTCCACAAACAGTGCCTTC 759
```

Search completed: August 1, 2003, 04:51:29
Job time : 237 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 1, 2003, 03:01:25 ; Search time 1455 Seconds
(without alignments)
2782.730 Million cell updates/sec

Title: US-10-010-408-2
Perfect score: 1440
Sequence: 1 MKGSPILHLATSLFLCLSM.....LCLPRPCILARSHSSWNSAF 250

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10010408/runat_25072003_101856_17934/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10010408.ecgn_1_1_1906_erunat_25072003_101856_17934 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1064	73.9	1058	14	BM805088 AGENCOURT
2	1043	72.4	979	14	BQ279131 AGENCOURT
3	977	67.8	886	13	BI822142 603039845
4	927.5	64.4	750	13	BM043988 603620978
5	918	63.7	1006	14	BM921531 AGENCOURT
6	911	63.3	651	13	BI457141 603185392
7	905	62.8	1073	14	BQ073722 AGENCOURT
8	902.5	62.7	928	13	BI161474 602864871
9	894	62.1	916	13	BI457367 603072631
10	858	59.6	888	13	BI825652 603072631
11	855.5	59.4	800	13	BI826781 603077268
12	815	56.6	790	13	BM046275 603626068
13	788	54.7	1022	14	BQ952960 AGENCOURT
14	785	54.5	1622	13	BG928868 HNC57-1-D
15	785	54.5	1166	13	BM543799 AGENCOURT
16	751.5	52.2	657	13	BM488499 pgm2n.pk0
17	721	50.1	1251	14	BQ961357 AGENCOURT
18	696	48.3	749	9	AL555144 AL555144
19	694	48.2	452	10	BE481184 166529 BA
20	668	46.4	933	14	BQ278961 AGENCOURT
21	648	45.0	489	14	BM751866 K-EST0028
22	540.5	37.5	618	12	BG538695 602566932
23	536	37.2	948	9	AL542806 AL542806
24	534	37.1	879	9	AL547439 AL547439
25	519	36.0	1002	9	AL556435 AL556435
26	498	34.6	1117	12	BF982720 602304922
27	496.5	34.5	2029	11	BC003205 Mus muscu
28	490.5	34.1	753	12	BE865071 UI-M-BH2
29	486	33.8	1057	12	BE898435 601681295
30	481.5	33.4	1898	11	AK013300 Mus muscu
31	473	32.8	940	14	BQ937887 AGENCOURT
32	467	32.4	405	12	BG900069 HOA51-1-A
33	463.5	32.2	909	12	BF981704 602306050
34	463.5	32.2	1076	13	BM554624 AGENCOURT
35	460	31.9	936	14	BQ880159 AGENCOURT
36	452.5	31.4	985	9	AL550072 AL550072
37	452	31.4	380	12	BG900020 HOA48-1-G
38	449	31.2	792	13	BI823598 603040962
39	444.5	30.9	1085	13	BM558306 AGENCOURT
40	444	30.8	693	13	BI824200 603040567
41	444	30.8	973	13	BM557871 AGENCOURT
42	442	30.7	826	12	BG116760 602318826
43	441.5	30.7	836	14	BQ716992 AGENCOURT
44	438.5	30.5	811	9	AU133455 AU133455
45	437	30.3	887	14	BQ894536 AGENCOURT

ALIGNMENTS

RESULT 1
BM805088
LOCUS BM805088 1058 bp mRNA linear EST 05-MAR-2002
DEFINITION AGENCOURT_6490429 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5725937
5' mRNA sequence.
ACCESSION BM805088
VERSION BM805088.1 GI:19121911
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1058)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12717 row: j column: 18
High quality sequence stop: 697.
Location/Qualifiers
1. 1058
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5725937"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcorV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcorV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."
BASE COUNT 139 a 423 c 294 g 200 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 1.67e-77 Length: 1058
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 14 Gaps: 0
US-10-010-408-2 (1-250) x BM805088 (1-1058)
QY 1 MetArgGlySerProLeuIleHisLeuValAlaThrSerPheLeuCysLeuSerMet 20
Db 9 ATGAGAGGACACCGAAGACCCACCTCTGCGCTTCTCCCTCCTGCTCTCAAG 68
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrProProGlnCysPro 40
Db 69 GTGGCGTACCAGCTGTGCCCCGACACGATGACTGCCCCCTGGCCACCTCCCGATGCCCG 128
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60
Db 129 CTGGGAGTACCCTGTGCTGTGATGGCTGTGGCTGCTGCCGGGTATGTGCACGGCGGCTG 188
QY 61 GlyGlySerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 189 GGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGGCGCTGTGCCAGGCC 248
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 249 GGGGAGACCCCGGTGACGGGGCGCTGTGCTCTTGGCAGAGGACGACAGCAGCTGT 308
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGlyLysThrPheLysProAsnCysArgValLeu 120
Db 309 GAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCACTGCAGCATCCGC 368
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuGlySerGluAspValArgLeu 140
Db 369 TGCCCGCTGCGAGGACGGCGGCTTACCTGCGCTGTGTGACGAGGATGTGCGGCTG 428
QY 141 ProSerTPAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 429 CCCAGCTGGGACTGCCCCACCCACGAGGAGGCTGAGGTCTGTGGCAAGTGCTGCCCTGAG 488

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 489 TGGGTGTGCGGCCCAAGAGGGGGAGCTGGGGACCCAGCCCTTCCAGCCCAAGAGCCAC 548
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThra 200
Db 549 TTTTCTGGCCTGTCTCTTCCCTGCCCCCTGTGTCTCCCTGCCCCAGATGAGCAGCGCC 608
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 609 TGGGAGCCCTGCTCGACACCACTGTGGCTGGGCATGGCCACCCGGGTGTCCAAACGAAAC 668
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 669 CGCTTGTGCGCACTGGAGACCCAGCGCGCTGTGCTGTCCAGGCGCCCTGCCACCCCTCC 728
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 729 AGGGGTGCGCACTCCACAAACAGTGCCTTC 758
RESULT 2
BQ279131 979 bp mRNA linear EST 07-MAY-2002
LOCUS AGENCOURT_7046721 NIH_MGC_107 Homo sapiens CDNA clone IMAGE:5805819
DEFINITION 5', mRNA sequence.
ACCESSION BQ279131 GI:20489339
VERSION BQ279131.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 979)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2045 row: k column: 04
High quality sequence stop: 752.
Location/Qualifiers
1. 979
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5805819"
/clone_lib="NIH_MGC_107"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pORF7; Site_1: EcorI;
Site_2: XhoI; CDNA made by oligo-dT priming.
Directionally cloned into EcorI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 141 a 347 c 313 g 170 t 8 others
ORIGIN
Alignment Scores:
Pred. No.: 8.03e-76 Length: 979
Score: 1043.00 Matches: 180
Percent Similarity: 80.66% Conservative: 16
Best Local Similarity: 74.07% Mismatches: 47
Query Match: 72.43% Indels: 0
DB: 14 Gaps: 0

US-10-010-408-2 (1-250) x BQ279131 (1-979)

QY	8	HisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMetValCysAlaGlnLeuCysArg	27
Db	14	CACTCTCTGGCCCTTCTCCCTCTCTGCTCTCTCAAAAGGTGCGTACCAGCTGTGCCCG	73
QY	28	ThrProCysThrCysProTrpThrProProGlnCysProGlnGlyValProLeuValLeu	47
Db	74	ACACCATGTACTCTGCCCTGCGCACCTCCCGATGCCCGCTGGGAGTACCCTGTGTCTG	133
QY	48	AspGlyCysGlyCysCysLysValCysAlaArgArgLeuGlyGluSerCysAspHisLeu	67
Db	134	GATGCTGTGGCTGTGCTGCGGGTATGTGCACGGCGGTGGGAGACCTTGCCAGCAACTC	193
QY	68	HisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHis	87
Db	194	CACGTCTGCGACGCCAGCCAGGGCGCTGTCTGACCAGCCCGGGGACAGACCCGGTGACGG	253
QY	88	GlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluValAsnGlyArgArgTyr	107
Db	254	GGGGCCCTGTGCTCTTGGCAGAGACGACAGACAGCTGTGAGTGAACGGCCCTGTAT	313
QY	108	LeuAspGlyGluThrPheLysProAsnCysArgValLeuCysArgCysAspAspGlyGly	127
Db	314	CGGGAAGGGGAGACCTTCCAGCCCACTGCAGCATCCGCTGCCCTGCCAGAGACGGCGGC	373
QY	128	PheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrpAspCysProArg	147
Db	374	TTCACCTGCGTGCCTGTGCGAGAGATGTGCGGCTGCCCACTGGGACTGCCCCAC	433
QY	148	ProLysArgIleGlnValProGlyLysCysCysProGluTrpValCysAspGlnGlyVal	167
Db	434	CCAGAGAGGGTGCAGGTCTGTGGCAAGTCTGCCCTGAGTGGGTGTGCGCAAGAGAGGG	493
QY	168	ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeuValThrPro	187
Db	494	GGACTGGGGAGACCCAGCCCTGCCAGCCCAAGAGACCCCACTTTCTGTGCCCTGTCTCC	553
QY	188	AlaSerAlaAspAlaProCysProAsnTrpSerThrAlaTrpGlyProCysSerThrThr	207
Db	554	CTGCCCCCTGGTGTCTCCCTGCCCAAGATGGAGACACGGCCCTGGGGACCCCTGACCAAC	613
QY	208	CysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsnArgPheCysGlnLeuGluIle	227
Db	614	TGTGGGCTGGGATGGCCACCCGGGTGTCCAACCAAGAACCCCTTCTGCGACTGGAGACC	673
QY	228	GlnArgArgLeuCysLeuProArgProCysLeuAlaAlaArgSerHisSerSerTrpAsn	247
Db	674	CAGCGCCGCTGTGCTGTCCAGGGCCCTGCCCAACCCCTCCAGGGGTGCGAGTCCACAACAC	733
QY	248	SerAlaPhe 250	
Db	734	AGTGCCTTC 742	

RESULT 3
LOCUS B1822142 886 bp mRNA linear EST 04-OCT-2001
DEFINITION 603039845F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180650 5',
mRNA sequence.
ACCESSION B1822142
VERSION B1822142
KEYWORDS B1822142.1 GI:159333692
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 886)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11450 row: j column: 11
High quality sequence stop: 865.
Location/Qualifiers
1. 886
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5180650"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 134 a 310 c 288 g 154 t
ORIGIN

Alignment Scores:
Pred. No.: 1.86e-70 Length: 886
Score: 977.00 Matches: 179
Percent Similarity: 78.71% Conservative: 17
Best Local Similarity: 71.89% Mismatches: 52
Query Match: 67.85% Indels: 3
DB: 13 Gaps: 0

US-10-010-408-2 (1-250) x B1822142 (1-886)

QY	3	GlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMetValCys	22
Db	1	GGCACACCGAAGACCCACCTCTGCTGCTTCTCCCTCTGCTCTCTCAAAAGTGGCT	60
QY	23	AlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysProGlnGly	42
Db	61	ACCAGCTGTGCCCGACACCAATGTACTGCCCCCTGGCCACTCCCGATGCCCGTGGGA	120
QY	43	ValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeuGlu	62
Db	121	GTACCCCTGTGCTGATGGCTGTGTGCTGTGCGGGTATGTGCACGGCGGTGGGGAG	180
QY	63	SerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnProGlyAla	82
Db	181	CCCTGGACCAACTCCACGTCCTGCCAGCCAGCCA-GGCTGTGTCGCCAGCCCGGGCA	239
QY	83	Gly-ProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluVa	102
Db	240	GGACCTGTGTGAGCGGGGGCCCTGTGCTCTTGGCAGAGACGACAGCAGCTGTGAGGT	299
QY	102	IAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeuCysAr	122
Db	300	GAAAGCGCCGCTGTATCGGGAAGGAGACCTTCCAGCCCACTGCAGCATCCGCTGCCG	359
QY	122	gCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeuProse	142
Db	360	CTGCGAGAGACGGCGGCTTCACTGCGCTGCGCGTGTGCAGAGATGTGCGGCTGCCAG	419
QY	142	rTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGluTrpVa	162
Db	420	CTGGACTGCCCCACCCCAAGAGGGTGCAGGTCTGGGCAAGTGTGCCCTGAGTGGGT	479
QY	162	lCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSe	182

Db	480	GTGGCGG-CAAGGAGGGGACTGGGGAGCCAGCCCTTCAGGCCCAAGGACCCAGTTTTC	538
Qy	182	ralaleuValThrProAlaSerAlaaspAlaProCysPheAsnTrpSerThrAlaTrpG1	202
Db	539	TGGCCTGTCTCTTCCTCCCTGCCCTGTGTCTCCCTGCCAGATGAGCAGCGCTGGGG	598
Qy	202	yProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsnArgph	222
Db	599	ACCCTGCTGCACCACCTGTGGGCTGGGCATGGCCACCGCGGTGTCCAACAGAACCGCTT	658
Qy	222	ecysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAlaArgse	242
Db	659	CTGCCGAACTGGAGACCAGCGCGGCTGTGCTGTCCAGGCGCTGCCACCTCCAGGGG	718
Qy	242	rHisSerSerTrpAsnSerAlaPhe250	
Db	719	TCGCAGTCCACAAACAGTGCCTTC	743
RESULT 4			
LOCUS	BM043988	750 bp	mrna
DEFINITION	603620978F1 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:5446794 5',		linear
ACCESSION	BM043988		
VERSION	BM043988.1	GI:16773255	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 750)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaphs-remail.nih.gov		
	Tissue Procurement: DCTD/DTP		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	plate: LLCM1929 row: k column: 19		
	High quality sequence stop: 714.		
FEATURES	Location/Qualifiers		
source	1. 750		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5446794"		
	/clone_1b="NIH_MGC_40"		
	/tissue_type="carcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;		
	Site_2: EcoRI; CDNA made by oligo-dT priming.		
	Directionally cloned into EcoRI/XhoI sites using the		
	following 5' adaptor: GGCACGAG(G). Library constructed by		
	Ling Hong in the laboratory of Gerald M. Rubin (University		
	of California, Berkeley) using ZAP-CDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies).		
	Note: this is a NIH_MGC Library.		
BASE COUNT	112 a 267 c 246 g 125 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.75e-66	Length:	750
Score:	927.50	Matches:	173
Percent Similarity:	79.83%	Conservative:	17
Best Local Similarity:	72.69%	Mismatches:	46
Query Match:	64.41%	Indels:	5
DB:	13	Gaps:	0

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	1	MetarglySerProleuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet	20	16	ATGAGAGGCACACCGAAGACCACCTCCTGGCCCTTCTCTCCTCTGCTCCTCAAG
BM921531	5', mRNA sequence.	BM921531	BM921531.1	GI:19371910	human.	21	ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro	40	76	GTCGCTACCCAGCTGTGCCCGACACACCATGATACCTGTGCCCTGGCCACCTCCCGATGCCG
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	41	GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu	60	136	CTGGAGTACCCCTGGTGTGATGATGCTGTGGCTGTGCCGGGTATGTGACAGCGGCTG
BM921531	5', mRNA sequence.	BM921531	BM921531.1	GI:19371910	human.	61	GlyGluSerCysAspHisLeuHisValCysAsp-ProSerGlnGlyLeuValCysGlnPr	80	196	GGGAGCCCTGGCACCACCTCCACGCTGTGCCAGCCCATGCCAGGGCCTGTGCCAGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	80	OgLYAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCy	100	256	CGGGCAGGACCCGGTGGACGGGGGGCCCTGTGCTCTTGGCAGAGGACGACAGCACTG
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	100	SgluValAsnGlyArgArgTyrLeuAspGlyGlyLthrPheLysProAsnCysArgValle	120	316	TGAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACTTCCAGCCCACTGCAGCATCCG
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	120	uCysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgle	140	376	CTGCCGCTGCAGAGACGGCGCTTACCTGTGCCCTGTGCAGCAGAGATGTGGGCT
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	140	uProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGl	160	436	GCCAGCTGGAGCTGCCCCACCCACGAGAGGTCGAGTCTTGGCAAGTGTGCCCTGA
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	160	uTrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThr-AlaGlnGlyHisG	180	496	GTGGGTGTGGCGCCACAGAGGGGGACTGGGGACCCAGCCCTTCAGCCCAAGACCCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	180	lnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThra	200	556	AGTTTCTGGCCCTGTCTCTTCCCTGCCCCCTGGTGTCCCTGCCCCAGATGAGACACGG
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	200	latrPglyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnA	220	616	CCTGGGACCCCTGTCTGAC-ACCTGTGGGCTGGGCAATGGCCACCCGGGTGTCCAACAGA
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	220	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM921531	BM921531.1	GI:19371910	Homo sapiens	236	snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro	236	675	ACCGTTCTAGCCGACTGGAGAC-CAGCGCCGCTG-TGCCCTGTCCAGGCC
BM921531	1006 bp mRNA	BM9215								


```
QY 33 ProTrpThrProProGlnCysProGlnGlyValProLeuValLeuAspGlyCysGlyCys 52
    ||||| |||||:::||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 CCTGGCCACCTCCCGATGCCGCTGGGAGTACCCCTGGTCTGGATGGCTGTGGCTGC 125
QY 53 CysLysValCysAlaArgArgLeuGlyGlySerCysAspHisLeuHisValCysAspPro 72
    ||:::||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TGCCGGGTATGTCACGGCGGCTGGGGAGCCCTGCGACCACTCCACGTCGCGACGCC 185
QY 73 SerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHisGlyAlaValCysLeu 92
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 AGCCA-GGCGTGTCTGCCAGCCCGGGGAGACCCGGTGGACGGGGGCCCTGTGCCCTC 244
QY 93 LeuAspGluAspAspGlySerCysGluValAsnGlyArgArgTyrLeuAspGlyGluThr 112
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 245 TTGGCAGAGAGACAGACAGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGAGACC 304
QY 113 PheLysProAsnCysArgValLeuCysArgCysAspAspGlyGlyPheThrCysLeuPro 132
    ||:::||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 305 TTCACAGCCCACTGCAGCATCCGCGTCCGCTGCGAGGAGCGCGGCTTCACCTGCGTCCG 364
QY 133 LeuCysSerGluAspValArgLeuProSerTrpAspCysProArgProLysArgIleGln 152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 365 CTGTGACAGCAGCATGTGCGGCTGCCAGCTGGGACTGCCCCACCCAGAGGGGTCCAG 424
QY 153 ValProGlyLysCysCysProGluTrpValCysAspGlnGlyValThrProAlaIleGln 172
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 425 GTCTGGGGCAAGTGTCTGCCCTGAGTGGGTGCGGCCAAGAGAGGGGAGTGGGAGACCAG 484
QY 173 ArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeuValThrProAlaSerAlaAspAla 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 485 CCCCTTCCAGCCCAAGAGACCCAGATTTCGTGCGCTTGTCTTCCCTGCCCTGTGTGT 544
QY 193 ProCysProAsnTrpSerThrAlaTrpGlyProCysSerThrThrCysGlyLeuGlyIle 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 545 CCCTGCCCAAGATGAGACAGCGGCTGGGGAGCCCTGTGACACACCTGTGGGCTGGGCATG 604
QY 213 AlaThrArgValSerAsnGlnAsnArgPheCysGlnLeuGlu 226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 605 GCCACCCGGGTGTCCAAGACAGAACCGGTCTGCGCAGCTGGAG 646

RESULT 7
BQ073722 1073 bp mRNA linear EST 02-APR-2002
LOCUS AGENCOURT_7046577 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:5806602
DEFINITION BQ073722 5' mRNA sequence.
ACCESSION BQ073722
VERSION BQ073722.1 GI:19902768
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1073)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2047 row: k column: 19
High quality sequence stop: 517.
Location/Qualifiers
FEATURES
source 1..1073
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5806602"
/clone_lib="NIH_MGC_101"
```

```
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)".
/notes="Organ: lung; Vector: pOTB7; Site_1: EcorI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcorI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC library."
```

BASE COUNT 166 a 384 c 347 g 176 t

ORIGIN

Alignment Scores:

Pred. No.:	1,91e-64	Length:	1073
Score:	905.00	Matches:	165
Percent Similarity:	77.35%	Conservative:	16
Best Local Similarity:	70.51%	Mismatches:	46
Query Match:	62.85%	Indels:	7
DB:	14	Gaps:	2

US-10-010-408-2 (1-250) x BQ073722 (1-1073)

```
QY 10 LeuAlaThrSerPheLeuCysLeuLeuSerMetValCysAlaGlnLeuCysArgThrPro 29
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 CTGGCCTTCTCCCTCTCTGCTGCTCTCAAGGTGCGTACCACGCTGTGCCGACACCA 61
QY 30 CysThrCysProTrpThrProProGlnCysProGlnGlyValProLeuValLeuAspGly 49
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 TGTACCTGCCCTGGCCACCTCCCGGATGCCGCTGGAGTACCCCTGTGCTGATGGC 121
QY 50 CysGlyCysCysLysValCysAlaArgArgLeuGlyGlySerCysAspHisLeuHisVal 69
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 TGTGGCTGTGCGGGGTATGTGCACGGCGCTGGGGAGCCCTGCCACCACTCCACGTC 181
QY 70 CysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHisGlyAla 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 TGGCAGCCAGCAGGGGCTGTGTGCCAGCCCGGGGAGACCCGCTGGCGGGGGGCC 241
QY 90 ValCysLeuLeuAspGluAspAspGlySerCysGluValAsnGlyArgArgTyrLeuAsp 109
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 CTGTGCTGTGGCAGAGAGACAGACAGCTGTGAGGTGAACGGCGCTGTATCGGGAA 301
QY 110 GlyGluThrPheLysProAsnCysArgValLeuCysArgCysAspAspGlyPheThr 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 GGGAGACCTTCCAGCCCACTGCAGCATCCGCTGCCGTGGAGAGCGCGCTTACC 361
QY 130 CysLeuProLeuCysSerGluAspValArgLeuProSerTrpAspCysProArgProLys 149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 TGCGTGCCTGTGCAGCAGGAGATGTGCGCTGCCACAGCTGGGACTGCCACCCACAGG 421
QY 150 ArgIleGlnValProGlyLysCysCysProGluTrpValCysAspGlnGlyValThrPro 169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 422 AGGCTCAGGCTCTGGCAAGTGTGCTGCTGAGTGGGTGCGGCCAAGAGGGGAGCTG 481
QY 170 AlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeuValThrProAlaSer 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 482 GGGACCCAGCCCTTCCAGCCCAAGGACCCAGTTTCTGTGCTGTCTTCCCTGCCCC 541
QY 190 AlaAspAlaProCysProAsnTrpSerThrAlaTrpGlyProCysSerThrThrCysGly 209
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 CCTGTGTCCCTGCCAGATGAGACAGGGCTGGGAGCCCTGTGACACACCTGTGGG 601
QY 210 Leu-GlyLeuAla--ThrArgValSerAsnGlnAsnArgPheCysGlnLeu-GluIleG 228
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 CTGGGGCATGGCCCAACCGGCTGTCCACCCAGAAACCGCTTTCGCCAGCTGGAGACC 661
QY 228 LnaArgArg-LeuCysLeu-----ProArgPro 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 662 AGCGCCGCTGTGCTGTGAGCAGGCGCCCTGGCCCC 697

RESULT 8
B1161474
```

LOCUS BI161474 928 bp mRNA linear EST 05-JUL-2001
DEFINITION 602864871F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5019065 5',
mRNA sequence.
ACCESSION BI161474
VERSION BI161474.1 GI:14621475
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 928)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
TITLE Email: cgapbs-remail.nih.gov
JOURNAL Tissue Procurement: ATCC
COMMENT CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM183 row: m column: 18
High quality sequence start: 28
High quality sequence stop: 756.
FEATURES
source
1. 928
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5019065"
/clone_lib="NIH_MGC_42"
/tissue_type="epitheloid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
BASE COUNT 161 a 301 c 309 g 157 t
ORIGIN
Alignment Scores:
Pred. No.: 2.56e-64 Length: 928
Score: 902.50 Matches: 172
Percent Similarity: 77.14% Conservative: 17
Best Local Similarity: 70.20% Mismatches: 49
Query Match: 62.67% Indels: 7
DB: 13 Gaps: 1
US-10-010-408-2 (1-250) x BI161474 (1-928)
QY 8 HisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMetValCysAlaGlnLeuCysArg 27
|||||
Db 83 CACCTCTGGCCTTCTCCCTCTCTGCTCTCTCAAAAGTGCCTACCCAGCTGTGCCCG 142
28 ThrProCysThrCysProTrpThrProProGlnCysProGlnGlyValProLeuValLeu 47
|||||
Db 143 ACACCATGTACCTGCCCCCTGGCCACCTCCCGCATGCCGCTGGGAGTACCCCTGGTCTG 202
48 AspGlyCysGlyCysCysLysValCysAlaArgArgLeuGlyLysCysAspHisLeu 67
|||||
Db 203 GATGGCTGTGGCTGCTGCCGGGTATGTGCACGGCGGCTGGGGAGAGCCCTGCACCACTC 262
68 HisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHis 87
|||||
Db 263 CACGTCTGCAGCGCCAGCCAGGGCCTGTGTGCCAGCCCGGGGAGACCGGCTGGACGG 322
88 GlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluValAsnGlyArgArgTyr 107

|||||
Db 323 GGGGCCCTGTGCTCTTGGCAGAGAGCAGACAGCTGTGAGGTGACAGCGCCCTGTAT 382
QY 108 LeuAspGlyGluThrPheLysProAsnGlyArgValLeuCysArgCysAspGlyGly 127
:::|||||
Db 383 CGGGAAGGGGAGACCTTCCAGCCCTGACGATCCGCTGCCCTGCCAGGAGCGCGCC 442
QY 128 PheThrCysLeuProLeuCysSerGlnAspValArgLeuProSerTrpAspCysProArg 147
|||||
Db 443 TTCACCTGCTGCTGCTGTGCAGCGAGATGTGCGCTGCCACGTGGAGCTGACCCCAA 502
QY 148 ProLysArgIleGlnValProGlyLysCysCysProGluTrpValCysAspGlnGlyVal 167
|||||
Db 503 CCCAGAGAGGCTCGAGGCTCTGCGCAAGTGTGCCCTGAGTGGGTGCGGCCCAAGAGGG 562
QY 168 ThrProAlaIleGlnArgSerThr-AlaGlnGlyHisGlnLeuSerAlaLeu-ValThrP 187
|||||
Db 563 GGACTGGGAGACCCAGACCTTACACGCCCAAGAGACCCAGATTCGTGCTCTTCC 622
QY 187 roAlaSerAlaAspAlaProCysProAsnTrpSerThrAlaTrpGly-ProCysSerThr 206
|||||
Db 623 CATGCCCCACTGTGTCCCTTGACCCAGATGAGACAGCGCTTGGGAGACCCCTGCTGACC 682
QY 207 ThrCys-GlyLeuGlyIleAla-ThrArgValSerAsnGlnAsnArgPheCysGlnLeu 226
|||||
Db 683 ACATGTGGGGCTGGCATTTGGACACCCGGGTGTCCCAACAGAACCGATTCTGCCGACTGG 742
QY 226 LuIleGln-ArgArgLeu--CysLeuProArgProCysLeuAlaAlaArgSerHisSer 244
|||||
Db 743 AGACCCCAAGCGCGCTAGTGGCTGTGTCACAGCCCTGCCCAACCTCCAGGGGTGCAAG 802
QY 245 Ser 245
|||
Db 803 TCC 805
RESULT 9
BI457367
LOCUS 916 bp mRNA linear EST 21-AUG-2001
DEFINITION 603185689F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5258398 5',
mRNA sequence.
ACCESSION BI457367
VERSION BI457367.1 GI:15248023
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 916)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
TITLE Email: cgapbs-remail.nih.gov
JOURNAL Tissue Procurement: ATCC
COMMENT CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1870 row: a column: 23
High quality sequence stop: 778.
FEATURES
source
1. 916
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5258398"
/clone_lib="NIH_MGC_42"
/tissue_type="epitheloid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the

Db 312 CGGAGAGGGAGACCTTCACGCCCCACATGACATCCGCTGCCGTCGAGAGGAGCGCGC 371
QY 128 PheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrpPaspCysProArg 147
Db 372 TTCACCTGCGTGCCGCTGTGCAGCGAGATGTGCGGCTGCCACGCTGGGACTGCCCCAC 431
QY 148 ProLysArgIleGlnValProGlyLysCys-CysProGluTrpValCysAspGlnGlyVal 167
Db 432 CCCAGGAGGGTCGAGGTCCTGTGGCAAGTCTTCCCTGAGTGGGTGGCGG-CAAGGAGG 490
QY 167 LThrProAlaIleGlnArgSerThrIleGlnGlyHisGlnLeuSerAlaLeuValThr-P 187
Db 491 GGAGCTGGGAGCCAGCCCTTCACGCCCCAGAGACCCAGTTTCTGGCCCTGTCTCTTC 550
QY 187 roAlaSerAlaAspAlaProCysProAsnTrpSerThrAlaTrpGlyProCysSerThr 207
Db 551 CCTGGCCCCCTGTGTGTCCTGCCAGATGGAGACAGGCTGT-GGACCCTGCTCGACCA 609
QY 207 hrcysglyleuglylleAlaThrArgValSerAsnGlnAsnArgPheCysGlnLeuGlu 227
Db 610 CCTGTGGGCTGGGCAT-GCCACCCGGGTGTCCAACAGAACCGTTCTGCCGACTGTGAGA 668
QY 227 leglnArgLeuCysLeuProArgProCysLeuAlaAlaArgSerHisSer-SerTrp 246
Db 669 CCCAGCGCGCTGTGCTGTCCAGGCGCTGCCACCCCTCCAGGGGTGGCAGTCCAAACA 728
QY 247 AsnSerAlaPhe 250
Db 729 AACAGTGCCTTC 740

RESULT 11
Bi826781 800 bp mRNA linear EST 04-OCT-2001
LOCUS 603077268F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169065 5',
DEFINITION mRNA sequence.
ACCESSION Bi826781
VERSION Bi826781.1 GI:15938331
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11420 row: 9 column: 18
High quality sequence stop: 788.
Location/Qualifiers
1. 800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5169065"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:

BASE COUNT 115 a 284 c 264 g 136 t 1 others
ALIGNMENT Scores:
Pred. No.: 1.54e-60 Length: 800
Score: 855.50 Matches: 170
Percent Similarity: 75.92% Conservative: 16
Best Local Similarity: 69.39% Mismatches: 56
Query Match: 59.41% Indels: 7
DB: 13 Gaps: 1

US-10-010-408-2 (1-250) x Bi826781 (1-800)

QY 8 HisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMetValCysAlaGlnLeuCysArg 27
Db 15 CACCTCCTGGCCTTCTCCTCCTCCTCCTCCTCAAGGTGCGTAACCCAGCTGTGCCCG 74
QY 28 ThrProCysThrCysProTrpThrProGlnCysProGlnGlyValProLeuValLeu 47
Db 75 ACACCATGTAACCTGCCCCCTGGCCACATCCCGATGCCCCGTGGGAGTAACCCCTGTGCTG 134
QY 48 AspGlyCysGlyCysCysLysValCysAlaArgArgLeuGlyGluSerCysAspHisLeu 67
Db 135 GATGGCTGTGGCTGTGCTGCCGATGTGTACAGGGCGCTGGGGAGCCCTGCGAACCACTC 194
QY 68 HisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHis 87
Db 195 CACGTCTGCAGCGCCAGCCA-GGCTGTCTGTGCCAGCCCGGGGAGACCCGGTGCCCG 253
QY 88 GlyAlaValCysLeuLeuAspGluAspGlySerCysGluValAsnGlyArgArgTyr 107
Db 254 GGGGCCCTGTGCTCTTGTGGCAGAGAGCACAGCAGCTGTGAGGTGAACGGCCGCTGTAT 313
QY 108 LeuAspGlyGluThrPheLysProAsnCysArgValLeuCysArgCysAspAspGlyGly 127
Db 314 CGGGAAGGGAGAACCTTCCAGCCCCACTGACATCCGCTGCCGTGCCGAGAGAGCGCGGC 373
QY 128 PheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrpPaspCysProArg 147
Db 374 TTCACCTGCGTGCCGCTGTGCAGCGAGATGTGCGGCTGCCACGCTGGGACTGCCCCAC 433
QY 148 ProLysArgIleGlnValProGlyLysCysCysProGluTrpValCysAspGlnGlyVal 167
Db 434 CCCAGGAGGGTCGAGGTCCTGTGGCAAGTCTGCCCTGAGTGGGTGGCGG-CAAGGAGGG 492
QY 168 ThrProAlaIleGlnArgSerThrIleGlnGlyHisGlnLeuSerAlaLeuValThr-Pr 187
Db 493 GGAGCTGGGAGCCAGCCCTTCACGCCCCAAGAGACCCAGTTTCTGGCCCTGTCTCTTC 552
QY 187 oAlaSerAlaAspAlaProCysProAsnTrpSerThrAlaTrpGlyProCysSerThrTh 207
Db 553 CTGCCCCCTGG--TGTCCCTGCCAGATGGAGNCACGGCTGGGA-CCCTGCTGACCCAC 608
QY 207 rcysglyleuglylleAlaThrArgValSerAsnGlnAsnArgPheCysGlnLeuGlu 227
Db 609 CTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACAGAACCG-TTCTGCCGACTGGAGAC 667
QY 227 eGlnArgLeuCysLeuProArgProCysLeuAlaAlaArgSerHisSerTrp-A 247
Db 668 CCAGCGCGCTGTGCTGTCCAGGCGCTGGCCACCCCTCCAGGGGTGGCAGTCCACAAAA 727
QY 247 snSerAlaPhe 250
Db 728 ACAGTGCCTTC 738

RESULT 12
BM046275 790 bp mRNA linear EST 07-NOV-2001
LOCUS 603626068F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5452480 5',
DEFINITION mRNA sequence.
ACCESSION BM046275
VERSION BM046275.1 GI:16775542

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1944 row: h column: 17
High quality sequence stop: 780.

FEATURES

source

1. 790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5452480"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

BASE COUNT 121 a 265 c 269 g 135 t
ORIGIN

Alignment Scores:

Pred. No.: 3.2e-57 Length: 790
Score: 815.00 Matches: 145
Percent Similarity: 78.82% Conservative: 15
Best Local Similarity: 71.43% Mismatches: 40
Query Match: 56.60% Indels: 4
DB: 13 Gaps: 1

US-10-010-408-2 (1-250) x BM046275 (1-790)

OY 51 GlyCysCysLysValCysAlaArgArgLeuGlyGluSerGysAspHisLeuHisValCys 70
|||||
Db 2 GGCTGCTGCCGGGTATGTGCACGGCGGCTGGGGAGCCCGCCGACCACTCCACGCTGC 61
OY 71 AspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlnHisGlyAlaVal 90
||| |||||||
Db 62 GACGCCAGCCAGGGCTGTCTGCCAGCCGGGGCAGAGCCCGGTGGCCGGGGCCCTG 121
OY 91 CysLeuLeuAsp-----GluAspAspGlySerCysGluValAsnGlyArgArgTyr 107
||||| |||||||
Db 122 TGCCCTCTGTAA-GCAGTGGCAGAGACGACGACGAGTGTGAGGTGACGCGCCCTGTAT 180
OY 108 LeuAspGlyGluThrPhelysProAsnGysArgValLeuGysArgCysAspAspGlyGly 127
:::|||||
Db 181 CGGAAGGGAGACCTTCCAGCCCACTGCACATCCGCGCCGCTGCAGAGACGGCGGC 240
OY 128 PheThrCysLeuProLeuCysSerGluAspValArgLeuProSerTrpAspCysProArg 147
|||||
Db 241 TTCACCTGCGTGGCGCTGTGCAGCGAGGATGTGGCGCTGCCACGCTGGAGCTGCCCCAC 300
OY 148 ProLysArgIleGlnValProGlyLysCysCysProGlyTrpValCysAspGlnGlyVal 167
|||:::|||||
Db 301 CCCAGGAGGGTGCAGGCTCTGTGGGCAAGTGTCTGAGTGGGTGTGGCGGCCAAGAGGG 360

OY 168 ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeuValThrPro 187
|||
Db 361 GGACTGGGAGACCCCTTCACAGCCCAAGGACCCCAAGTTTCTGTGCTCTCTCC 420
OY 188 AlaSerAlaAspAlaProCysProAsnTrpSerThrAlaTrpGlyProCysSerThrThr 207
|||||
Db 421 CTGCCCCCTGTGTCCCTGCCCAAGATGAGACACGGCTGGGAGACCTGTGCACACAC 480
OY 208 CysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsnArgPheCysGlnLeuGluIle 227
|||||
Db 481 TGTGGCTGGCATGCGCCACCCGGGTGTCCACCAAGAACCGCTTGTGCGACTGGAGACC 540
OY 228 GlnArgArgLeuCysLeuProArgProCysLeuAlaAlaArgSerHisSerSerTrpAsn 247
|||||
Db 541 CAGCGCCGCTGTGCTGTCCAGGCCCTGCCACACCTCCAGGGGTGCGAGTCCACAAC 600
OY 248 SerAlaPhe 250
Db 601 AGTGCCCTTC 609

RESULT 13

BQ952960 1022 bp mRNA linear EST 21-AUG-2002
LOCUS BQ952960
DEFINITION AGENCOURT_8931684 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6484123
5', mRNA sequence.
ACCESSION BQ952960
VERSION BQ952960.1 GI:22368438
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1022)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2669 row: m column: 20
High quality sequence stop: 482.

FEATURES

source

1. 1022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6484123"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

BASE COUNT 129 a 305 c 457 g 131 t
ORIGIN

Alignment Scores:
Pred. No.: 7.22e-55 Length: 1022
Score: 788.00 Matches: 139
Percent Similarity: 79.79% Conservative: 15
Best Local Similarity: 72.02% Mismatches: 36
Query Match: 54.72% Indels: 4
DB: 14 Gaps: 1

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12361 row: f column: 15
High quality sequence stop: 547.

FEATURES

source

Location/Qualifiers
1. 1166
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="NIH_MGC_125"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT 216 a 424 c 315 g 207 t 4 others
ORIGIN

Alignment Scores:

Pred. No.: 1.5e-54 Length: 1166
Score: 785.00 Matches: 149
Percent Similarity: 73.54% Conservative: 15
Best Local Similarity: 66.82% Mismatches: 43
Query Match: 54.51% Indels: 16
DB: 13 Gaps: 6

US-10-010-408-2 (1-250) x BM543799 (1-1166)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20
DB 211 ATGAGAGGCACACCGAAGACCCTCTGCGCTTCTGCTCTCTCTCTCAAG 270
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrProProGlnCysPro 40
DB 271 GTGCGTACCAGCTGTGCGCCGACACATGTACTGCGCTGCGCCACCTCCCGATGCCG 330
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60
DB 331 CTGGAGTACCCCTGTGCTGTGATGTCTGTGCTGTGCTGCGGTATGTGCACGGCGCTG 390
QY 61 GlyIleSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
DB 391 GGGAGCCCTGCGACCACTCCACTCTGCGACGCCACGAGGGCCTGTGCTGCCAGCCC 450
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
DB 451 GGGGACGAGACCCGGTGGACGGGGGGCTGTGCTCTTGGCAGAGGACGACAGCAGCTGT 510
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGlyIleThrPheIleProAsnCysArgValLeu 120
DB 511 GAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCACTGCAGCATCCGC 570
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
DB 571 TGCCGCTGCGAGACGGCGCTTCACTGTGCGCTGTGCAAGCAAGATGTGCGGCTG 630
QY 141 ProSerTrpAspCysProArg-ProLysArgIleGlnValProGlyLysCysCysProGln 160
DB 631 CCCAGCTGGGACTGCCCCACCCCAAGAGAGGTGCAAGTCCCTGNGCAAGTGTGCCCTGA 690
QY 160 uTrpValCysAspGln-----GlyValThrProAlaIleGlnArgSerTh 175

DB 691 ATGGGTGTGCGGCCCAAGAGAGGGGAGCTGGGGGACCAACCC-----TTCCA 738
QY 175 rAlaGlnGlyHisGln--LeuSerAlaLeuValThrPro--AlaSerAlaAspAlaPr 193
DB 739 ACCCCAGGAGACCCAGNNTTCTGTGGGCTGTCTCTTCCCTGGCCCTGGNGGTCCCC 798
QY 193 oCys-ProAsnTrp--SerThrAlaTrpGlyProCysSerThrThrCys-----GlyL 210
DB 799 CTGCCCAAAATGGGAGCAACGGGCTGGGGGACCCCTGTCTCAACAACCTGTGGGGCC 858
QY 210 euGly 211
DB 859 TGGGC 863

Search completed: August 1, 2003, 04:47:29
Job time : 1466 secs

New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
Claim 33; Page 181-182; 284pp; English.

The present invention describes Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2 and WISP-3 have homology to connective tissue growth factor (CTGF). Products from the present invention can be used to treat WISP-related disorders such as breast, ovarian, and colon cancer or melanoma. The products can be used to treat arteriosclerosis. The products can also be used to treat other diseases e.g. benign and malignant tumours, leukaemia and lymphoid malignancies, neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal, and blastocoeleic disorders, haematopolesis-related disorders, tissue-growth disorders, skin disorders, desmoplasia, fibrotic lesions, kidney disorders, bone-related disorders such as osteoporosis, trauma such as burns, incisions, and other wounds, connective tissue disorders, catabolic states, testicular-related disorders, and inflammatory, angiogenic and immunologic disorders including arteriosclerosis. The products can also be used for detection and diagnosis especially of individuals with neoplastic cell growth or proliferation. The products can be used in the production of transgenic or knock-out animals. Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing cells.

XX Sequence 251 AA;

Query Match 90.9%; Score 1308.5; DB 20; Length 251;
Best Local Similarity 90.0%; Pred. No. 1e-90;
Matches 226; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

OY 1 MRGSPILHLATSFLLSMVCAQLCRPTCTCPWTPQCPGVPVLVDGCGCKVCARRL 60
Db 1 MRGNPLHLAISFLCILSMVYSQLCPAPCAPWTPPQCPGVPVLVDGCGCKVCARRL 60
OY 61 GESCDHLHVCDFPSQGLVCPGAGPGGHCAYCLLDEDDGSCCEVNGRRYLDGETFKPNCRVL 120
Db 61 GESCDHLHVCDFPSQGLVCPGAGPGGHCAYCLLDEDDGSCCEVNGRRYLDGETFKPNCRVL 120
OY 121 CRCDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIQRSTAQGH 179
Db 121 CRCDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIQRSTAQGH 180
OY 180 QLSALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRPCLAA 239
Db 181 QLSALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLSRPCLAA 240
OY 240 ARSHSSWNSAF 250
Db 241 SRSHGSWNSAF 251

RESULT 4
ID AAY17680 standard; Protein; 250 AA.

XX AAY17680;

DT 06-AUG-1999 (first entry)

DE Mouse WISP-2 protein SEQ ID NO:78.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopolesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder.

OS Mus sp.

XX

PN WO9921998-A1.

XX 06-MAY-1999.

XX 29-OCT-1998; 98WO-US22991.

XX 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

XX (GENTH) GENENTECH INC.

XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

XX WPI; 1999-337420/28.

XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

XX Disclosure; Page 245-246; 284pp; English.

CC The present invention describes Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2 and WISP-3 have homology to connective tissue growth factor (CTGF). Products from the present invention can be used to treat WISP-related disorders such as breast, ovarian, and colon cancer or melanoma. The products can be used to treat arteriosclerosis. The products can also be used to treat other diseases e.g. benign and malignant tumours, leukaemia and lymphoid malignancies, neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal, and blastocoeleic disorders, haematopolesis-related disorders, tissue-growth disorders, skin disorders, desmoplasia, fibrotic lesions, kidney disorders, bone-related disorders such as osteoporosis, trauma such as burns, incisions, and other wounds, connective tissue disorders, catabolic states, testicular-related disorders, and inflammatory, angiogenic and immunologic disorders including arteriosclerosis. The products can also be used for detection and diagnosis especially of individuals with neoplastic cell growth or proliferation. The products can be used in the production of transgenic or knock-out animals. Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing cells.

XX Sequence 250 AA;

Query Match 90.5%; Score 1303.5; DB 20; Length 250;
Best Local Similarity 90.0%; Pred. No. 2.5e-90;
Matches 225; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

OY 2 RGSPLIHLATSFLLSMVCAQLCRPTCTCPWTPPQCPGVPVLVDGCGCKVCARRLG 61

Db 1 RGNPLHLAISFLCILSMVYSQLCPAPCAPWTPPQCPGVPVLVDGCGCKVCARRLG 60

OY 62 ESCDHLHVCDFPSQGLVCPGAGPGGHCAYCLLDEDDGSCCEVNGRRYLDGETFKPNCRVLC 121

Db 61 ESCDHLHVCDFPSQGLVCPGAGPGGHCAYCLLDEDDGSCCEVNGRRYLDGETFKPNCRVLC 120

OY 122 RCDDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIQRSTAQGHQ 180

Db 121 RCDDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIQRSTAQGHQ 180

OY 181 LLSALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLSRPCLAA 240

Db 181 LLSALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLSRPCLAA 240

OY 241 RSHSSWNSAF 250

Db 241 RSHGSWNSAF 250

RESULT 5
ID AAY17681

XX AAY17681 standard; Protein; 249 AA.

XX

XX Claim 24; Fig 1; 108pp; English.
PS

This represents a rat heparin-induced CCN-like protein (HICP) protein. Agents that stimulate or inhibit HICP protein activity or expression, antisense HICP nucleic acid molecules and HICP antibodies, can be used to modulate cell-associated activity. HICP modulators can be used to treat disorders characterized by aberrant HICP protein activity or expression. Probes capable of hybridizing to HICP mRNA or antibodies specific for HICP can be used to detect HICP activity in a biological sample. HICP can be used to treat disorders, such as a cardiovascular or fibrotic disorder, characterized by aberrant cell proliferation.

SQ Sequence 250 AA;

Query Match	100.0%;	Score 1440;	DB 20;	Length 250;
Best Local Similarity	100.0%;	Pred. No. 1.4e-100;		
Matches 250; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MRGSPLIHLATSFLLCLSMVCAQLCRTPTCTCPWTPPQCPOGVPLVLDGCGCKVCARRL	60
Db	1	MRGSPLIHLATSFLLCLSMVCAQLCRTPTCTCPWTPPQCPOGVPLVLDGCGCKVCARRL	60
QY	61	GESCDHLVCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCENVNGRRYLDGETFKPNCRVL	120
Db	61	GESCDHLVCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCENVNGRRYLDGETFKPNCRVL	120
QY	121	CRCDDBGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVTPAIQRSTAQHQ	180
Db	121	CRCDDBGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVTPAIQRSTAQHQ	180
QY	181	LSALVTPASADAPCPNMWSTAMGPCSTTCGLGIATRVSNQNRFCQLEIQRRLLCPRPCLAA	240
Db	181	LSALVTPASADAPCPNMWSTAMGPCSTTCGLGIATRVSNQNRFCQLEIQRRLLCPRPCLAA	240
QY	241	RSHSSWNSAF 250	
Db	241	RSHSSWNSAF 250	

RESULT 2
AAV27440
ID AAV27440 standard; Protein: 227 AA.

AAV27440;
26-NOV-1999 (first entry)
Rat HICP mature polypeptide.
Heparin-induced CCN-like protein; HICP; cell-associated activity;
cardiovascular disorder; aberrant cell proliferation; fibrotic disorder
Rattus sp.
WO9947556-A2.
23-SEP-1999.
18-MAR-1999; 99WO-US05999.
19-MAR-1998; 98US-0044273.
(TUFT) TUFTS COLLEGE.
Castellot JJ;

; 1999-562060/47.
SDB; AAZ07521.

leic acid sequences encoding rat heparin-induced CCN-like protein in methods to identify modulators or in diagnostic applications

PS Disclosure; Fig 2; 108pp; English.

CC The invention provides a rat heparin-induced CCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents the rat HICP mature polypeptide.

...	
50	Sequence 227 AA;

Query Match	91.9%;	Score 1323;	DB 20;	Length 227;
Best Local Similarity	100.0%;	Pred. No. 7.7e-92;		
Matches 227; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	24	QLCRPTCTCPWTPPQCPQGVPLVLDDGCGCKVCARRLGESCDHLHVCDPSQGLVQCQPGAG	83
Db	1	QLCRPTCTCPWTPPQCPQGVPLVLDDGCGCKVCARRLGESCDHLHVCDPSQGLVQCQPGAG	60
QY	84	PGHGAVCLLDEDDGSCVEVNGRRYLDGETFKPNCRVLCRCDDGGFTCLPLCSEDEVRLPSW	143
Db	61	PGHGAVCLLDEDDGSCVEVNGRRYLDGETFKPNCRVLCRCDDGGFTCLPLCSEDEVRLPSW	120
QY	144	DCPRPRKRIQVPGKCCPEWVCDQGYVTPAIQIRSTAQGHQLSALVTPASADAPCPNMWSTAMCP	203
Db	121	DCPRPRKRIQVPGKCCPEWVCDQGYVTPAIQIRSTAQGHQLSALVTPASADAPCPNMWSTAMCP	180
QY	204	CSTTCGLGIATRVSNQNRFCQLEIQRRCLPRPCLAAARSHSSWNSAF	250
Db	181	CSTTCGLGIATRVSNQNRFCQLEIQRRCLPRPCLAAARSHSSWNSAF	227

RESULT 3
AAV17651
ID AAV17651 standard; Protein; 251 AA

	AAV17651;	06-AUG-1999 (first entry)	Mouse WISP-2 protein SEQ ID NO:20.
XX			
XX			
DT			
DE			
XX			
KW			
KW			
KW			
KW			
KW			
KW			
KW			
KW			
XX			
OS			
XX			
PN			
PN			
XX			
PD			
PF			
XX			
XX			
PR			
PR			
PR			

PI	Botstein DA,	Cohen RL,	Goddard A,	Gurney AL,	Hillan K,
PI	Lawrence DA,	Levine AJ,	Pennica D,	Roy MA,	Wood WI,
XX					
DR	WPL; 1999-337420/28.				
DR	N-PSDB; AAX76488.				

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

protein - protein search, using sw model

on: July 25, 2003, 12:25:25 ; Search time 84 Seconds
(without alignments)
472.401 Million cell updates/sec

Title: US-10-010-408-2

Perfect score: 1440

Sequence: 1 MRGSPHLHLATSEFLCLSM.....LCLEPRCLARSHSSWNSAF 250

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1440	100.0	250	20	AAV27434
2	1323	91.9	227	20	AAV27440
3	1308.5	90.9	251	20	AAV17651
4	1303.5	90.5	250	20	AAV17680
5	1298.5	90.2	249	20	AAV17681
6	1292.5	89.8	248	20	AAV17682
7	1291.5	89.7	247	20	AAV17683
8	1284.5	89.2	246	20	AAV17684
9	1280.5	88.9	245	20	AAV17685

10	1276.5	88.6	244	20	AAV17686	Mouse WISP-2 prote
11	1268.5	88.1	243	20	AAV17687	Mouse WISP-2 prote
12	1264.5	87.8	242	20	AAV17688	Mouse WISP-2 prote
13	1260.5	87.5	241	20	AAV17689	Mouse WISP-2 prote
14	1257.5	87.3	239	20	AAV17691	Mouse WISP-2 prote
15	1257.5	87.3	240	20	AAV17690	Mouse WISP-2 prote
16	1253.5	87.0	238	20	AAV17692	Mouse WISP-2 prote
17	1247.5	86.6	237	20	AAV17693	Mouse WISP-2 prote
18	1243.5	86.4	236	20	AAV17694	Mouse WISP-2 prote
19	1234.5	85.7	235	20	AAV17695	Mouse WISP-2 prote
20	1232.5	85.6	234	20	AAV17696	Mouse WISP-2 prote
21	1228.5	85.3	233	20	AAV17697	Mouse WISP-2 prote
22	1224.5	85.0	232	20	AAV17698	Mouse WISP-2 prote
23	1219.5	84.7	231	20	AAV17699	Mouse WISP-2 prote
24	1217.5	84.5	229	20	AAV17701	Mouse WISP-2 prote
25	1217.5	84.5	230	20	AAV17700	Mouse WISP-2 prote
26	1216.5	84.5	228	20	AAV17650	Mouse putative mat
27	1064	73.9	250	19	AAW37946	Human connective t
28	1064	73.9	250	20	AAV17649	Human WISP-2 prote
29	1064	73.9	250	20	AAV05285	EGF-like homologue
30	1064	73.9	250	20	AAW94616	Human growth facto
31	1064	73.9	250	21	AAAB24402	Human PRO261 prote
32	1064	73.9	250	21	AAAB8573	Human PRO261 amino
33	1064	73.9	250	22	AAU12331	Human PRO261 polyp
34	1064	73.9	250	22	AAAB85526	Human secreted pro
35	1064	73.9	250	22	AAAB85542	Human secreted pro
36	1064	73.9	250	22	AAAB84599	Connective tissue
37	1064	73.9	250	22	AAAB68598	PRO261. Homo sapi
38	1064	73.9	250	22	AAAB53084	Human angiogenesis
39	1064	73.9	250	23	ABG32502	Human wild-type WI
40	1064	73.9	250	23	ABG76572	Human secreted pro
41	1064	73.9	250	23	ABG76588	Human secreted pro
42	1064	73.9	250	23	ABG64716	Human albumin fusi
43	1064	73.9	250	23	ABG64717	Human albumin fusi
44	1064	73.9	250	24	ABU66729	Human PRO polypept
45	1064	73.9	250	24	ABU67005	Human secreted/ttra

ALIGNMENTS

RESULT 1	
AAV27434	standard; Protein; 250 AA.
XX	
AC	AAV27434;
XX	
DT	26-NOV-1999 (first entry)
XX	
DE	Rat HICP polypeptide.
XX	
KW	Heparin-induced CCN-like protein; HICP; cell-associated activity; rat;
KW	cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX	
OS	Rattus sp.
XX	
PN	WO9947556-A2.
XX	
PD	23-SEP-1999.
XX	
PF	18-MAR-1999; 99WO-US05999.
XX	
PR	19-MAR-1998; 98US-0044273.
XX	
PA	(TUFT) TUFTS COLLEGE.
XX	
PI	Castellot JJ;
XX	
DR	WPI; 1999-562060/47.
DR	N-PSDB; AAZ07516, AAZ07517.
XX	
PT	Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT	used in methods to identify modulators or in diagnostic applications

XX	Sequence	248 AA;	89.8%;	Score 1292.5;	DB 20;	Length 248;
XX	Query Match		89.8%;	Pred. No. 1.6e-89;		
XX	Best Local Similarity		89.8%;	Pred. No. 1.6e-89;		
XX	Matches 223;	Conservative	9;	Mismatches 15;	Indels 1;	Gaps 1
QY	4 SPLIHLATSFLLSMVCAQLCRTPTCTCPWTPPQCPQGVPLVLDGCGCKVCARRLIGES	63				
DB	1 NPLIHLAISFLCILSMVYSQILCPAPACPCPWTTPQCPGVPLVLDGCGCKVCARRLIGES	60				
QY	64 CDHLHVCDPSQGLVCPGAGPGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRVLCRC	123				
DB	61 CDHLHVCDPSQGLVCPGAGPGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRVLCRC	120				
QY	124 DDGFTCLPLCSEDRVRLPSWDCPRPKRIQVPKGCCPEWVCDQGV-TPAIQRTAOGHQLS	182				
DB	121 DDGFTCLPLCSEDRVRLPSWDCPRPKRIQVPKGCCPEWVCDQGV-TPAIQRTAOGHQLS	180				
QY	183 ALVTPASADAPCPNMWSTAWGFCSTTCGLGIATRVSNQNRFCQLEIQRRLCLPRPCLARS	242				
DB	181 ALVTPASADAPCPNMWSTAWGFCSTTCGLGIATRVSNQNRFCQLEIQRRLCLSRPCLARS	240				
QY	243 HSSWNSAF 250					
DB	241 HSSWNSAF 248					
RESULT 7						
AA17683						
ID	AA17683 standard; Protein; 247 AA.					
XX	AA17683;					
XX	06-AUG-1999 (first entry)					
DT	Mouse WISP-2 protein SEQ ID NO:81.					
DE						
XX	WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;					
KW	connective tissue growth factor; cancer; melanoma; arteriosclerosis;					
KW	leukaemia; lymphoid malignancy; haematopoiesis-related disorder;					
KW	tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;					
KW	kidney disorder; bone-related disorder; osteoporosis; trauma; burn;					
KW	connective tissue disorder; catabolic state; inflammation;					
KW	testicular-related disorder; angiogenesis; immunological disorder.					
OS	Mus sp.					
XX						
PN	W09921998-A1.					
XX						
PD	06-MAY-1999.					
XX						
PF	29-OCT-1998; 98WO-US22991.					
XX						
PR	14-APR-1998; 98US-0081695.					
PR	29-OCT-1997; 97US-0063704.					
XX	03-FEB-1998; 98US-0073612.					
PA	(GETH) GENENTECH INC.					
XX						
PI	Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;					
PI	Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;					
XX						
DR	WPI; 1999-337420/28.					
XX						
PT	New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3					
XX						
PS	Disclosure; Page 248; 284pp; English.					
XX						
CC	The present invention describes Wnt-1 induced secreted polypeptides,					
CC	WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2					
CC	and WISP-3 have homology to connective tissue growth factor (CTGF).					
CC	Products from the present invention can be used to treat WISP-related					

Query	Match	Best Local Similarity	Score	DB	Length
Matches	223;	Conservative	89.7%;	90.3%;	Pred. No. 1.9e-89;
			8;	Mismatches	15;
				Indels	1;
				Gaps	1
CC	disorders such as breast, ovarian, and colon cancer or melanoma. The				
CC	products can be used to treat arteriosclerosis. The products can also be				
CC	used to treat other diseases e.g. benign and malignant tumours,				
CC	leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,				
CC	hypothalamic and other glandular, macrophagal, epithelial, stromal, and				
CC	blastocoeleic disorders, haematopoesis-related disorders, tissue-growth				
CC	disorders, skin disorders, desmoplasia, fibrotic lesions, kidney				
CC	disorders, bone-related disorders such as osteoporosis, trauma such as				
CC	burns, incisions, and other wounds, connective tissue disorders,				
CC	catabolic states, testicular-related disorders, and inflammatory,				
CC	angiogenic and immunologic disorders including arteriosclerosis. The				
CC	products can also be used for detection and diagnosis especially of				
CC	individuals with neoplastic cell growth or proliferation. The products				
CC	can be used in the production of transgenic or knock-out animals.				
CC	Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing				
CC	cells.				
XX					
XX	Sequence	247 AA;			
QY	Query Match	89.7%;	Score 1291.5;	DB 20;	Length 247;
Db	1	PLIHLATSFCLLSMVCQAQLCRPTCTCPWTPPQCQGVPLVLDGCGCKVCARRLGESC	64		
QY	125	DGFTCLPLCSEEDVRLPSWDCPRBRKRIQVPGKCCPEWVCDQGV-TPAIQRSTAQGHQLSA	183		
Db	121	DGFTCLPLCSEEDVRLPSWDCPRBRKRIQVPGRCCEPWVCDQAVMPAIPQSSAQGHQLSA	180		
QY	184	LVTASADAPCPNMSTAMGPCSTTCGLGATATRVSNQNRFCQLEIQRRLCLPRPCLARSH	243		
Db	181	LVTASADAPCPNMSTAMGPCSTTCGLGATATRVSNQNRFCQLEIQRRLCLSRPCLARSH	240		
QY	244	SSWNSAF 250			
Db	241	GSWNSAF 247			
RESULT 8					
AA17684					
ID	AA17684	standard; Protein; 246 AA.			
XX	AA17684;				
DT	06-AUG-1999	(first entry)			
DE	Mouse WISP-2 protein SEQ ID NO:82.				
XX					
XX	WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;				
XX	connective tissue growth factor; cancer; melanoma; arteriosclerosis;				
XX	leukaemia; lymphoid malignancy; haematopoesis-related disorder;				
XX	tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;				
XX	kidney disorder; bone-related disorder; osteoporosis; trauma; burn;				
XX	connective tissue disorder; catabolic state; inflammation;				
XX	testicular-related disorder; angiogenesis; immunological disorder.				
XX	Mus sp.				
XX					
XX	WO9921998-A1.				
XX					
XX	06-MAY-1999.				
XX					
XX	29-OCT-1998;	98WO-US22991.			
XX					
XX	14-APR-1998;	98US-0081695.			
XX	29-OCT-1997;	97US-0063704.			
XX	03-FEB-1998;	98US-0073612.			

XX (GETH) GENENTECH INC.
PA Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
XX Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
PI WPI; 1999-337420/28.
XX
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
XX Disclosure; Page 249; 284pp; English.
XX
XX The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 246 AA;

Query Match 89.2%; Score 1284.5; DB 20; Length 246;
Best Local Similarity 90.2%; Pred. No. 6.5e-89;
Matches 222; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY 6 LIHLATSFLLCLSMVCAQLCRTPTCTCPWTPPQCQGVPLVLDGCGCKVCARLGESCD 65
DB 1 LIHLAISFLCILSMVYSQLCRAPACAPWTPPQCQGVPLVLDGCGCKVCARLGESCD 60
QY 66 HLHVCDPSQGLVQCQPGAGSGHGAVALDDEDDGSCVENGRRLDGETFEKPNCRVLCRCD 125
DB 61 HLHVCDPSQGLVQCQPGAGSGHGAVALDDEDDGSCVENGRRLDGETFEKPNCRVLCRCD 120
QY 126 GFTCLPLCSEDEVRLPSWDCPRPRRIQVPGKCCPEWVCDQGV-TPAIQSTAGHQLSAL 184
DB 121 GFTCLPLCSEDEVRLPSWDCPRPRRIQVPGKCCPEWVCDQAVMPAIQPSAOGHQLSAL 180
QY 185 VTPASADAPCPNMSTAWGPCSTTCGLGIAITRVSNQNRFCOLEIQRLCLPRPCLARSHS 244
DB 181 VTPASADGCPNMSTAWGPCSTTCGLGIAITRVSNQNRFCOLEIQRLCLSRPCLASRSHG 240
QY 245 SWNSAF 250
DB 241 SWNSAF 246

RESULT 9
AAV17685
ID AAV17685 standard; Protein; 245 AA.
XX
AC AAV17685;
XX
DT 06-AUG-1999 (first entry)
XX
DE Mouse WISP-2 protein SEQ ID NO:83.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;

KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder.
XX
OS Mus sp.
XX
XX WO9921998-A1.
PN
XX
XX 06-MAY-1999.
PD
XX
XX 29-OCT-1998; 98WO-US22991.
PF
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
XX WPI; 1999-337420/28.
DR
XX
XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
PT
XX
PS Disclosure; Page 249-250; 284pp; English.
XX
XX The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 245 AA;

Query Match 88.9%; Score 1280.5; DB 20; Length 245;
Best Local Similarity 90.2%; Pred. No. 1.3e-88;
Matches 221; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY 7 IHLATSFLLCLSMVCAQLCRTPTCTCPWTPPQCQGVPLVLDGCGCKVCARLGESCDH 66
DB 1 IHLAISFLCILSMVYSQLCRAPACAPWTPPQCQGVPLVLDGCGCKVCARLGESCDH 60
QY 67 LHVCDPSQGLVQCQPGAGSGHGAVALDDEDDGSCVENGRRLDGETFEKPNCRVLCRCD 126
DB 61 LHVCDPSQGLVQCQPGAGSGHGAVALDDEDDGSCVENGRRLDGETFEKPNCRVLCRCD 120
QY 127 GFTCLPLCSEDEVRLPSWDCPRPRRIQVPGKCCPEWVCDQGV-TPAIQSTAGHQLSALV 185
DB 121 GFTCLPLCSEDEVRLPSWDCPRPRRIQVPGKCCPEWVCDQAVMPAIQPSAOGHQLSALV 180
QY 186 TPAADAPCPNMSTAWGPCSTTCGLGIAITRVSNQNRFCOLEIQRLCLPRPCLARSHS 245
DB 181 TPAADGCPNMSTAWGPCSTTCGLGIAITRVSNQNRFCOLEIQRLCLSRPCLASRSHG 240
QY 246 WNSAF 250
DB 111111

PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
XX WPI; 1999-337420/28.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Disclosure; Page 253-254; 284pp; English.
XX
XX The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 241 AA;

Query Match 87.5%; Score 1260.5; DB 20; Length 241;
Best Local Similarity 90.0%; Pred. No. 4e-87;
Matches 217; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY 11 ATSELCILSMVCAOLCCTPCTCPMTTPQCPQGVPLVLDGCGCKVCARRLGESCDHLHVC 70
Db 1 AISFLCILSMVYSQLCRAPACAPMTTPQCPQGVPLVLDGCGCCRCVCAARRLGESCDHLHVC 60

QY 71 DPSQGLVCQPGAGPGSGHGAVALLDDEDDGSCSEVNGRRYLDGETFKPNCRYLRCDDGGFTC 130
Db 61 DPSQGLVCQPGAGPGSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRYLRCDDGGFTC 120

QY 131 LPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIQRSTAGHQLSALVTPAS 189
Db 121 LPLCSEDEVRLPSWDCPRPKRIQVPGRCCEWVCDQAVMQPAIQSSAQGHQLSALVTPAS 180

QY 190 ADAPCPNMWSTAWGPGSTTCGLGIATRVSNQNRFCQLEIQRLCLPRPCLARSHSSWNSA 249
Db 181 ADGCPNMWSTAWGPGSTTCGLGIATRVSNQNRFCQLEIQRLCLSRPCLARSHSGSWNSA 240

QY 250 F 250
Db 241 F 241

RESULT 14
ID AAY17691 standard; Protein; 239 AA.
XX
AC AAY17691;
XX

DT 06-AUG-1999 (first entry)
XX
DE Mouse WISP-2 protein SEQ ID NO:89.
XX
XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder.
XX
OS Mus sp.
XX
PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
XX (GETH) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
XX WPI; 1999-337420/28.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Disclosure; Page 255-256; 284pp; English.
XX
XX The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 239 AA;

Query Match 87.3%; Score 1257.5; DB 20; Length 239;
Best Local Similarity 90.4%; Pred. No. 6.7e-87;
Matches 216; Conservative 8; Mismatches 14; Indels 1; Gaps 1;

QY 13 SFLLCLSMVCAOLCCTPCTCPMTTPQCPQGVPLVLDGCGCKVCARRLGESCDHLHVCDP 72
Db 1 SFLLCLSMVYSQLCRAPACAPMTTPQCPQGVPLVLDGCGCCRCVCAARRLGESCDHLHVCDP 60

QY 73 SQGLVCQPGAGPGSGHGAVALLDDEDDGSCSEVNGRRYLDGETFKPNCRYLRCDDGGFTCLP 132
Db 61 SQGLVCQPGAGPGSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRYLRCDDGGFTCLP 120

QY 133 LCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIQRSTAGHQLSALVTPASAD 191
Db 121 LCSEDEVRLPSWDCPRPKRIQVPGRCCEWVCDQAVMQPAIQSSAQGHQLSALVTPASAD 180

QY 192 APCPNWSTAWGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRPCLARSHSSWNSAF 250
DB 181 GPCPNWSTAWGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLSRPCLASRSHGWSNSAF 239

RESULT 15

AAV17690
ID AAV17690 standard; Protein; 240 AA.

AC AAV17690;

DT 06-AUG-1999 (first entry)

DE Mouse WISP-2 protein SEQ ID NO:88.

KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
connective tissue growth factor; cancer; melanoma; arteriosclerosis;
leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
connective tissue disorder; catabolic state; inflammation;
testicular-related disorder; angiogenesis; immunological disorder.

OS Mus sp.

PN WO9921998-A1.

PD 06-MAY-1999.

PF 29-OCT-1998; 98WO-US22991.

PR 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

PA (GETH) GENENTECH INC.

PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

DR WPI; 1999-337420/28.

PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

PS Disclosure; Page 254-255; 284pp; English.

CC The present invention describes Wnt-1 induced secreted polypeptides,
WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
used to treat other diseases e.g. benign and malignant tumours,
leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
hypothalamic and other glandular, macrophagal, epithelial, stromal, and
blastocoele disorders, haematopoiesis-related disorders, tissue-growth
disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
disorders, bone-related disorders such as osteoporosis, trauma such as
burns, incisions, and other wounds, connective tissue disorders,
catabolic states, testicular-related disorders, and inflammatory,
angioenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
cells.

SQ Sequence 240 AA;

Query Match 87.3%; Score 1257.5; DB 20; Length 240;
Best Local Similarity 90.4%; Pred. No. 6.7e-87;
Matches 216; Conservative 8; Mismatches 14; Indels 1; Gaps 1;

QY 13 SFLCLLSMWCAQLCRTPTCTCPWTPPQCPQGVPLVLDGCGCKVCARRLGESCDHLHVCDP 72
DB 2 SFLGILSMVYSQLCRAPACAPWTPPQCPGVPLVLDGCGCCRCVRCARLGESCDHLHVCDP 61
QY 73 SQGLVCQPGAGPGGCGAVCLLDEDDGSCFVNGRRYLDGETFKPNCRVLCRDDGGFTCLP 132
DB 62 SQGLVCQPGAGPGGCGAVCLFEEDDGSCFVNGRRYLDGETFKPNCRVLCRDDGGFTCLP 121
QY 133 LCSEDRVLPSPWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIQRSTAOQHLSALVTPASAD 191
DB 122 LCSEDRVLPSPWDCPRPKRIQVPGKCCPEWVCDQAVMOPAIQSSAQGHLSALVTPASAD 181
QY 192 APCPNWSTAWGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRPCLARSHSSWNSAF 250
DB 182 GPCPNWSTAWGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLSRPCLASRSHGWSNSAF 240

Search completed: July 25, 2003, 12:34:59
Job time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:33:35 ; Search time 29 Seconds
(without alignments)
364.749 Million cell updates/sec

Title: US-10-010-408-2
Perfect score: 1440
Sequence: 1 MRGSPILHLATSLFLCLSM.....LCLPRPCLARSHSSWNSAF 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308.5	90.9	251	4	US-09-182-145-20 Sequence 20, Appl
2	1303.5	90.5	250	4	US-09-182-145-78 Sequence 78, Appl
3	1298.5	90.2	249	4	US-09-182-145-79 Sequence 79, Appl
4	1292.5	89.8	248	4	US-09-182-145-80 Sequence 80, Appl
5	1291.5	89.7	247	4	US-09-182-145-81 Sequence 81, Appl
6	1284.5	89.2	246	4	US-09-182-145-82 Sequence 82, Appl
7	1280.5	88.9	245	4	US-09-182-145-83 Sequence 83, Appl
8	1276.5	88.6	244	4	US-09-182-145-84 Sequence 84, Appl
9	1268.5	88.1	243	4	US-09-182-145-85 Sequence 85, Appl
10	1264.5	87.8	242	4	US-09-182-145-86 Sequence 86, Appl
11	1260.5	87.5	241	4	US-09-182-145-87 Sequence 87, Appl
12	1257.5	87.3	239	4	US-09-182-145-89 Sequence 89, Appl
13	1257.5	87.3	240	4	US-09-182-145-88 Sequence 88, Appl
14	1253.5	87.0	238	4	US-09-182-145-90 Sequence 90, Appl
15	1247.5	86.6	237	4	US-09-182-145-91 Sequence 91, Appl
16	1243.5	86.4	236	4	US-09-182-145-92 Sequence 92, Appl
17	1234.5	85.7	235	4	US-09-182-145-93 Sequence 93, Appl
18	1232.5	85.6	234	4	US-09-182-145-94 Sequence 94, Appl
19	1228.5	85.3	233	4	US-09-182-145-95 Sequence 95, Appl
20	1224.5	85.0	232	4	US-09-182-145-96 Sequence 96, Appl
21	1219.5	84.7	231	4	US-09-182-145-97 Sequence 97, Appl
22	1217.5	84.5	229	4	US-09-182-145-99 Sequence 99, Appl
23	1217.5	84.5	230	4	US-09-182-145-98 Sequence 98, Appl
24	1216.5	84.5	228	4	US-09-182-145-19 Sequence 19, Appl
25	1064	73.9	250	4	US-09-182-145-16 Sequence 16, Appl
26	1059	73.5	249	4	US-09-182-145-56 Sequence 56, Appl
27	1054	73.2	248	4	US-09-182-145-57 Sequence 57, Appl

28	1048	72.8	247	4	US-09-182-145-58 Sequence 58, Appl
29	1047	72.7	246	4	US-09-182-145-59 Sequence 59, Appl
30	1043	72.4	243	4	US-09-182-145-62 Sequence 62, Appl
31	1043	72.4	244	4	US-09-182-145-61 Sequence 61, Appl
32	1043	72.4	245	4	US-09-182-145-60 Sequence 60, Appl
33	1035	71.9	242	4	US-09-182-145-63 Sequence 63, Appl
34	1031	71.6	241	4	US-09-182-145-64 Sequence 64, Appl
35	1027	71.3	240	4	US-09-182-145-65 Sequence 65, Appl
36	1025	71.2	238	4	US-09-182-145-67 Sequence 67, Appl
37	1025	71.2	239	4	US-09-182-145-66 Sequence 66, Appl
38	1021	70.9	236	4	US-09-182-145-69 Sequence 69, Appl
39	1021	70.9	237	4	US-09-182-145-68 Sequence 68, Appl
40	1017	70.6	235	4	US-09-182-145-70 Sequence 70, Appl
41	1008	70.0	234	4	US-09-182-145-71 Sequence 71, Appl
42	1004	69.7	233	4	US-09-182-145-72 Sequence 72, Appl
43	1000	69.4	232	4	US-09-182-145-73 Sequence 73, Appl
44	997	69.2	230	4	US-09-182-145-75 Sequence 75, Appl
45	997	69.2	231	4	US-09-182-145-74 Sequence 74, Appl

ALIGNMENTS

RESULT 1
US-09-182-145-20
; Sequence 20, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 20
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-182-145-20

Query Match 90.9%; Score 1308.5; DB 4; Length 251;
Best local Similarity 90.0%; Pred. No. 1.1e-105;
Matches 226; Conservative 9; Mismatches 15; Indels 1; Gaps 1;
QY 1 MRGSPILHLATSLFLCLSMVCAQLCTPCTCPWTTPQCPQGVPLVLDGCGCKVCARRL 60
DB 1 MRGNPLHLAISFLCILSMVYSQLCPACPCPWTTPQCPGVPLVLDGCGCCRCVARRL 60
QY 61 GESCDHLHVCPSQGLVCQPGAGPGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRYL 120
DB 61 GESCDHLHVCPSQGLVCQPGAGPGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRYL 120
QY 121 CRCDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIORSTAQGH 179
DB 121 CRCDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIORSTAQGH 179
QY 121 CRCDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIORSTAQGH 179
DB 121 CRCDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIORSTAQGH 179
QY 180 QLSALVTPASADAPCPNMWSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRLCLPRPCLA 239
DB 180 QLSALVTPASADAPCPNMWSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRLCLPRPCLA 239

DB	QY	DB
181 QLSALVTPASADGPCPNWSTAMGPCSTTCGLGIATRYNSNQNRFCQLEIQRLCLSRPCLA	240 ARSHSSWNSAF	240
	:	
	241 SRSHGSWNSAF	251

RESULT 2

```

US-09-182-145-78
; Sequence 78, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Bolstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 78
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-182-145-78

```

Query Match	90.5%	Score 1303.5;	DB 4;	Length 250;
Best Local Similarity	90.0%	Pred. No. 2.9e-105;		
Matches 225; Conservative	9;	Mismatches 15;	Indels 1;	Gaps 1;

[illegible]

RESULT 3

US-09-182-145-79
; Sequence 79, Application US/09182145B
; Patent NO. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.

```

: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182,145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 79
: LENGTH: 249
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-182-145-79

```

Query Match	90.28;	Score 1298.5;	DB 4;	Length 249;
Best Local Similarity	90.08;	Pred. No. 7.8e-105;		
Matches 224; Conservative	9;	Mismatches 15;	Indels 1;	Gaps 1;

[illegible]

RESULT 4

```

US-09-182-145-80
; Sequence 80, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14

```


RESULT 7

US-09-182-145-83
 ; Sequence 83, Application US/09182145B
 ; Patent No. 6387657
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David A.
 ; APPLICANT: Cohen, Robert
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Lawrence, David A.
 ; APPLICANT: Levine, Arnold J.
 ; APPLICANT: Pennica, Diane
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: P1176R2
 ; CURRENT APPLICATION NUMBER: US/09/182,145B
 ; CURRENT FILING DATE: 1998-10-29
 ; EARLIER APPLICATION NUMBER: US 60/063,704
 ; EARLIER FILING DATE: 1997-10-29
 ; EARLIER APPLICATION NUMBER: US 60/073,612
 ; EARLIER FILING DATE: 1998-02-04
 ; EARLIER APPLICATION NUMBER: US 60/081,695
 ; EARLIER FILING DATE: 1998-04-14
 ; NUMBER OF SEQ ID NOS: 156
 ; SEQ ID NO 83
 ; LENGTH: 245
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-182-145-83

Query Match 88.9%; Score 1280.5; DB 4; Length 245;
 Best Local Similarity 90.2%; Pred. No. 2.8e-103;
 Matches 221; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY 7 IHLATSFILCLISWCAQLCRPTCTPWTTPQCPQGVPLVLDGCGCKVCARRLGESCDH 66
 1 IHLAISFLCILSMWYSQLCFAPACACPTTPQCPGVPVLVDGCGCKVCARRLGESCDH 60
 QY 67 LHVCDPSQGLVCQPGAGPGGHGAVCLDEDDGSCSEVNGRRYLDGETFKPNCRVLCRCDDG 126
 61 LHVCDPSQGLVCQPGAGPGGHGAVCLDEDDGSCSEVNGRRYLDGETFKPNCRVLCRCDDG 120
 QY 127 GFTCLPLCSEEDVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIORSTAQGHQLSALV 185
 121 GFTCLPLCSEEDVRLPSWDCPRPKRIQVPGKCCPEWVCDQAVMOPAIQPSAOGHQLSALV 180
 QY 186 TPASADAPCPNWSWAGPCSTTCGLGATRVSNQNRFCQLEIQRRLCLPRCLARSHSS 245
 181 TPASADGCPNWSWAGPCSTTCGLGATRVSNQNRFCQLEIQRRLCLPRCLARSHSS 240
 QY 246 WNSAF 250
 241 WNSAF 245

RESULT 8
 US-09-182-145-84
 ; Sequence 84, Application US/09182145B
 ; Patent No. 6387657
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David A.
 ; APPLICANT: Cohen, Robert
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Lawrence, David A.
 ; APPLICANT: Levine, Arnold J.
 ; APPLICANT: Pennica, Diane
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: P1176R2
 ; CURRENT APPLICATION NUMBER: US/09/182,145B
 ; CURRENT FILING DATE: 1998-10-29
 ; EARLIER APPLICATION NUMBER: US 60/063,704
 ; EARLIER FILING DATE: 1997-10-29
 ; EARLIER APPLICATION NUMBER: US 60/073,612
 ; EARLIER FILING DATE: 1998-02-04
 ; EARLIER APPLICATION NUMBER: US 60/081,695
 ; EARLIER FILING DATE: 1998-04-14
 ; NUMBER OF SEQ ID NOS: 156
 ; SEQ ID NO 84
 ; LENGTH: 244
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-182-145-84

Query Match 88.6%; Score 1276.5; DB 4; Length 244;
 Best Local Similarity 90.2%; Pred. No. 6.1e-103;
 Matches 220; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY 8 HLLATSFILCLISWCAQLCRPTCTPWTTPQCPQGVPLVLDGCGCKVCARRLGESCDH 67
 1 HLLAISFLCILSMWYSQLCFAPACACPTTPQCPGVPVLVDGCGCKVCARRLGESCDH 60
 QY 68 HVCDPQGLVCQPGAGPGGHGAVCLDEDDGSCSEVNGRRYLDGETFKPNCRVLCRCDDG 127
 61 HVCDPQGLVCQPGAGPGGHGAVCLDEDDGSCSEVNGRRYLDGETFKPNCRVLCRCDDG 120
 QY 128 FTCLPLCSEEDVRLPSWDCPRPKRIQVPGKCCPEWVCDQGV-TPAIORSTAQGHQLSALV 186
 121 FTCLPLCSEEDVRLPSWDCPRPKRIQVPGKCCPEWVCDQAVMOPAIQPSAOGHQLSALV 180
 QY 187 PASADAPCPNWSWAGPCSTTCGLGATRVSNQNRFCQLEIQRRLCLPRCLARSHSSW 246
 181 PASADGCPNWSWAGPCSTTCGLGATRVSNQNRFCQLEIQRRLCLPRCLARSHSSW 240
 QY 247 NSAF 250
 241 NSAF 244

RESULT 9
 US-09-182-145-85
 ; Sequence 85, Application US/09182145B
 ; Patent No. 6387657
 ; GENERAL INFORMATION:
 ; APPLICANT: Botstein, David A.
 ; APPLICANT: Cohen, Robert
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Lawrence, David A.
 ; APPLICANT: Levine, Arnold J.
 ; APPLICANT: Pennica, Diane
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: P1176R2
 ; CURRENT APPLICATION NUMBER: US/09/182,145B
 ; CURRENT FILING DATE: 1998-10-29
 ; EARLIER APPLICATION NUMBER: US 60/063,704
 ; EARLIER FILING DATE: 1997-10-29
 ; EARLIER APPLICATION NUMBER: US 60/073,612
 ; EARLIER FILING DATE: 1998-02-04
 ; EARLIER APPLICATION NUMBER: US 60/081,695
 ; EARLIER FILING DATE: 1998-04-14
 ; NUMBER OF SEQ ID NOS: 156
 ; SEQ ID NO 85
 ; LENGTH: 243
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-182-145-85

	Query Match	88.1%;	Score 1268.5;	DB 4;	Length 243;	
	Best Local Similarity	90.1%;	Pred. No. 3e-102;			
	Matches 219;	Conservative	8;	Mismatches 15;	Indels 1;	Gaps 1
QY	9 LLATSFLLLSMVCQAOLCRTPCTCPWTPPQCQPQGVPLVLDDGCGCKKVCARRLGESCDHLH	: : :	68			
Db	1 LLAISFLCILSMVYSQLCPAPCACPWTPQCPCPGVPLVLDDGCGCCRCVARRLGESCDHLH		60			
QY	69 VCDPSOGLVCQPGAGPGGHGAVCLLDEDDGSCEVNNGRRYLDETFFKPNCRYLCRCDDDGEF	:	128			
Db	61 VCDPSOGLVCQPGAGPSGRGAVCLFEEDDGSCEVNNGRRYLDGETFFKPNCRYLCRCDDDGEF		120			
QY	129 TCLPLCSEDEVRLPSWDCPRPKRIQVP GKCCPEWVCDQGV-TPAIQNSTAQGHQLSALVTP	: : :	187			
Db	121 TCLPLCSEDEVRLPSWDCPRPRRIQVPGRCCEPEWVCDQAVVMQPAIQSPSSAOGHQLSALVTP		180			
QY	188 ASADAPCPNMSTAWGPCSTTCGLGIATRVSQNÖRFQLEIQRRILCLPRCLLARSHSSWN		247			
Db	181 ASADGPCPNMSTAWGPCSTTCGLGIATRVSQNÖRFQLEIQRRILCLSRPCLASRSRHSWN		240			
QY	248 SAF 250					
Db	241 SAF 243					

```

RESULT 10
US-09-182-145-86
: Sequence 86, Application US/09182145B
: Patent No. 6387657
: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182,145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 86
: LENGTH: 242
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-182-145-86

```

	Query Match	87.8%;	Score 1264.5;	DB 4;	Length 242;
	Best Local Similarity	90.1%;	Pred. No. 6.6e-102;		
	Matches 218;	Conservative 8;	Mismatches 15;	Indels 1;	Gaps 1;
QY	10 LATSFLCLSMVCAOLCRTPTCTCPMTTPPQCQGVPVLVDGGCCCKVCARLGE	SCDHLHV	69		
Dd	1 LAISFICILSMVYSQLCPAFCACPMWTPPQCPSGVPLVDGCGCCRCVCARLGE	SCDHLHV	60		
QY	70 CDPSQGLVCQPAGAPSGHGAVCLDEDGDGSCEVNGRRYLDGETFKPNCRVL	CRCDDGGFT	129		
Dd	61 CDPSSQGLVCQPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRVL	CRCDDGGFT	120		
QY	130 CLPLCSEDRVLRPSWDGPRPKRIQVPGKCCPEWVCDOGV-TPAIORSTAQGH	LSALVTPA	188		
Dd	121 CLPLCSEDRVLRPSWDGPRPRIQVPGRCCEPWVCDAQVAIQPSSAQGHLSAL	VTPA	180		

[illegible]

```

RESULT 11
US-09-182-145-87
; Sequence 87, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 87
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-182-145-87

```

```
Query Match      87.5%; Score 1260.5; DB 4; Length 241;
Best Local Similarity 90.0%; Pred. No. 1.5e-101;
Matches 217; Conservative 8; Mismatches 15; Indels 1; Gaps 1;
```

OY 11 ATSFLLLSMYCAQLCRTPTCTCPWTTPQCPOGVPLVLDGCGCKVCARRLGESCDHLHVC 70
 | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db 1 AISFLCILSMYSQLCPPACPCACPMTTPQCPRPGVPVLVDGCGCCRCVCCARRLGESCDHLHVC 60

OY 71 DPSOGLVCOPGAPEGGHGAVCLLDEDDGSCEVNRRYLDETETKPNCRVLCRCDGGEFTC 130
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 DPSOGLVCOPGAEPGRGAVALCFEEDDGSCEVNRRYLDGETFKPNCRVLCRCDDGGEFTC 120

OY 131 LPLCSSEDEVRLPSPMDCPRPKRIQVP GKCCPEWVCDQGV--TPAIQRSTAQGHQLSALVTPAS 189
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 LPLCSSEDEVRLPSPMDCPRPRIQVP GRCCPEWVCDQA VMPAIQPSAQGHQLSALVTPAS 180

OY 190 ADAPCPNMSTAMGPCSTTTCGLGIATRVSNQNRFQLEIQRRLCLPRPCLARSHSSWN SA 249
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 ADGPCPMNSTAMGPCSTTTCGLGIATRVSNNRNFCOLEIORRLCLSRPCLARS HSGSWNSA 240

OY 250 F 250
 |
Db 241 F 241

RESULT 12
US-09-182-145-89
; Sequence 89, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert


```
Qy      193 PCPNWSTAWGPCSTTCGLGIATRVSNQNRFQQLIEIQRRLCLPRPCLARSHSSWN$AF   250  
        |||||  
Db     181 PCPNWSTAWGPCSTTCGLGIATRVSNQNRFQGLEIQRRLLCLRPPCLASRSHGSWN$AF   238
```

RESULT 15

```

US-09-182-145-91
; Sequence 91, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 91
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-182-145-91

```

Query Match	86.6%;	Score 1247.5;	DB 4;	Length 237;
Best Local Similarity	90.3%;	Pred. No. 1.9e-100;		
Matches 214; Conservative	8;	Mismatches 14;	Indels 1;	Gaps 1;

QY	15	LCILSMVCAQLCRTPTCTCPMTTPPOCGVPLVLDGCGCKVCARRLGESCDHLHVCDPSQ	74
		: :	
Db	1	LCILSMVYSQLCPAPCACPMWTPPOCGVPLVLDGCGCCRVCARRLGESCDHLHVCDPSQ	60
QY	75	GLVCQPGAGPGGGHGAVALDDEDGSCFEVNGRRYLDGETFKPNCRYLCRCDDGGFTCLPLC	134
		:	
Db	61	GLVCQPGAGPSGRGAVCLFEEDDGSCFEVNGRRYLDGETFKPNCRYLCRCDDGGFTCLPLC	120
QY	135	SEDVRLPSWDCPRPKRIQVPKGKCCPEWVCDQGV-TPAIQRSTAQGHQLSALVTPASADAP	1933
		: : :	
Db	121	SEDVRLPSWDCPRPRRIQVPGRCCPEWVCDQAVMQPAIQSSAQGHQLSALVTPASADGP	180
QY	194	CPNWSTAWGPCSTTCGLGIATRVSNQNFQLEIQRLCLPRPCLARSHSSWNSAF	250
Db	181	CPNWSTAWGPCSTTCGLGIATRVSNQNFQLEIQRLCLSRPCLARSHSGWNSAF	237

Search completed: July 25, 2003, 12:38:36
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:35:05 ; Search time 23 Seconds
(without alignments)
1290.867 Million cell updates/sec

Title: US-10-010-408-2
Perfect score: 1440
Sequence: 1 MRGSPLIHLATSFCLLSM.....LCLPRPCLARSHSSWNSAF 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	100.0	250	14 US-10-010-408-2	Sequence 2, Appl1
2	1323	91.9	227	14 US-10-010-408-13	Sequence 13, Appl1
3	1308.5	90.9	251	15 US-10-112-267-20	Sequence 20, Appl1
4	1303.5	90.5	250	15 US-10-112-267-78	Sequence 78, Appl1
5	1298.5	90.2	249	15 US-10-112-267-79	Sequence 79, Appl1
6	1292.5	89.8	248	15 US-10-112-267-80	Sequence 80, Appl1
7	1291.5	89.7	247	15 US-10-112-267-81	Sequence 81, Appl1
8	1284.5	89.2	246	15 US-10-112-267-82	Sequence 82, Appl1
9	1280.5	88.9	245	15 US-10-112-267-83	Sequence 83, Appl1
10	1276.5	88.6	244	15 US-10-112-267-84	Sequence 84, Appl1
11	1268.5	88.1	243	15 US-10-112-267-85	Sequence 85, Appl1
12	1264.5	87.8	242	15 US-10-112-267-86	Sequence 86, Appl1
13	1260.5	87.5	241	15 US-10-112-267-87	Sequence 87, Appl1
14	1257.5	87.3	239	15 US-10-112-267-89	Sequence 89, Appl1
15	1257.5	87.3	240	15 US-10-112-267-88	Sequence 88, Appl1

16	1253.5	87.0	238	15 US-10-112-267-90	Sequence 90, Appl1
17	1247.5	86.6	237	15 US-10-112-267-91	Sequence 91, Appl1
18	1243.5	86.4	236	15 US-10-112-267-92	Sequence 92, Appl1
19	1234.5	85.7	235	15 US-10-112-267-93	Sequence 93, Appl1
20	1232.5	85.6	234	15 US-10-112-267-94	Sequence 94, Appl1
21	1228.5	85.3	233	15 US-10-112-267-95	Sequence 95, Appl1
22	1224.5	85.0	232	15 US-10-112-267-96	Sequence 96, Appl1
23	1219.5	84.7	231	15 US-10-112-267-97	Sequence 97, Appl1
24	1217.5	84.5	229	15 US-10-112-267-99	Sequence 99, Appl1
25	1217.5	84.5	230	15 US-10-112-267-98	Sequence 98, Appl1
26	1216.5	84.5	228	15 US-10-112-267-19	Sequence 19, Appl1
27	1064	73.9	250	10 US-09-915-582-53	Sequence 69, Appl1
28	1064	73.9	250	10 US-09-915-582-69	Sequence 320, App
29	1064	73.9	250	15 US-10-028-072-320	Sequence 320, App
30	1064	73.9	250	15 US-10-121-049-320	Sequence 320, App
31	1064	73.9	250	15 US-10-123-904-320	Sequence 320, App
32	1064	73.9	250	15 US-10-140-470-320	Sequence 320, App
33	1064	73.9	250	15 US-10-175-746-320	Sequence 320, App
34	1064	73.9	250	15 US-10-176-918-320	Sequence 320, App
35	1064	73.9	250	15 US-10-176-921-320	Sequence 320, App
36	1064	73.9	250	15 US-10-137-865-320	Sequence 320, App
37	1064	73.9	250	15 US-10-140-474-320	Sequence 320, App
38	1064	73.9	250	15 US-10-142-431-320	Sequence 320, App
39	1064	73.9	250	15 US-10-143-114-320	Sequence 320, App
40	1064	73.9	250	15 US-10-140-002-320	Sequence 320, App
41	1064	73.9	250	15 US-10-142-419-320	Sequence 320, App
42	1064	73.9	250	15 US-10-123-262-320	Sequence 320, App
43	1064	73.9	250	15 US-10-142-423-320	Sequence 320, App
44	1064	73.9	250	15 US-10-121-050-320	Sequence 320, App
45	1064	73.9	250	15 US-10-141-755-320	Sequence 320, App

ALIGNMENTS

RESULT 1
US-10-010-408-2
Sequence 2, Application US/10010408
Publication No. US20020165185A1

GENERAL INFORMATION:
APPLICANT: John J. Castelliott, Jr.
TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules and Uses Therefor

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MB1-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-010-408-2

Query Match 100.0%; Score 1440; DB 14; Length 250;
Best Local Similarity 100.0%; Pred. No. 6.9e-103;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGSPILHLATSLFLCLSMVCAQLCRTPCTPMTPPQCPGVPLVLDGCGCKVCARRL 60
|||||
Db 1 MRGSPILHLATSLFLCLSMVCAQLCRTPCTPMTPPQCPGVPLVLDGCGCKVCARRL 60
QY 61 GESCDHLHVCDPSSQGLVCPGAGPGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRYL 120
|||||
Db 61 GESCDHLHVCDPSSQGLVCPGAGPGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRYL 120
QY 121 CRCDDGFTCLPLCSEDEVRLPSMDCPRPKRIQVPKCCPEWVCDQGVTPAIQSTAGHQ 180
|||||
Db 121 CRCDDGFTCLPLCSEDEVRLPSMDCPRPKRIQVPKCCPEWVCDQGVTPAIQSTAGHQ 180
QY 181 LSAVTPASADAPCPNMSTAMGPGSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRCLAA 240
|||||
Db 181 LSAVTPASADAPCPNMSTAMGPGSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRCLAA 240
QY 241 RSHSSWNSAF 250
|||||
Db 241 RSHSSWNSAF 250

RESULT 2

US-10-010-408-13
Sequence 13, Application US/10010408
Publication No. US20020165185A1

GENERAL INFORMATION:

APPLICANT: John J. Castellot, Jr.
TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules
and Uses Therefor

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/010,408

FILING DATE: 07-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/044,273

FILING DATE: March 19, 1998

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: MBI-004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 227 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-010-408-13

Query Match 91.9%; Score 1323; DB 14; Length 227;
Best Local Similarity 100.0%; Pred. No. 3.8e-103;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QLCRPTCTPMTPPQCPGVPLVLDGCGCKVCARRLGESCDHLHVCDPSSQGLVCPGAG 83
|||||
Db 1 QLCRPTCTPMTPPQCPGVPLVLDGCGCKVCARRLGESCDHLHVCDPSSQGLVCPGAG 60
QY 84 PGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRYLCRDDGGFTCLPLCSEDEVRLPSW 143
|||||
Db 61 PGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRYLCRDDGGFTCLPLCSEDEVRLPSW 120
QY 144 DCPRRKRIQVPKCCPEWVCDQGVTPAIQSTAGHQLSALVTPASADAPCPNMSTAMGP 203
|||||
Db 121 DCPRRKRIQVPKCCPEWVCDQGVTPAIQSTAGHQLSALVTPASADAPCPNMSTAMGP 180
QY 204 CSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRCLAAARSHSSWNSAF 250
|||||
Db 181 CSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRCLAAARSHSSWNSAF 227

RESULT 3

US-10-112-267-20

Sequence 20, Application US/10112267

Publication No. US20030068678A1

GENERAL INFORMATION:

APPLICANT: Botstein, David A.

APPLICANT: Cohen, Robert

APPLICANT: Goddard, Audrey

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Lawrence, David A.

APPLICANT: Levine, Arnold J.

APPLICANT: Pennica, Diane

APPLICANT: Roy, Margaret Ann

APPLICANT: Wood, William I.

TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: P1176R2

CURRENT APPLICATION NUMBER: US/10/112,267

CURRENT FILING DATE: 2002-03-27

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B

PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704

PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612

PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695

PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14

NUMBER OF SEQ ID NOS: 156

SEQ ID NO 20

LENGTH: 251

TYPE: PRT

ORGANISM: Mus musculus

US-10-112-267-20

Query Match 90.9%; Score 1308.5; DB 15; Length 251;
Best Local Similarity 90.0%; Pred. No. 6.8e-102;
Matches 226; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 1 MRGSPILHLATSLFLCLSMVCAQLCRTPCTPMTPPQCPGVPLVLDGCGCKVCARRL 60
|||||
Db 1 MRGSPILHLATSLFLCLSMVCAQLCRTPCTPMTPPQCPGVPLVLDGCGCKVCARRL 60
QY 61 GESCDHLHVCDPSSQGLVCPGAGPGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRYL 120
|||||
Db 61 GESCDHLHVCDPSSQGLVCPGAGPGHGAVALDDEDDGSCVNGRRYLDGETFKPNCRYL 120


```
QY 121 CRDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDOGV-TPAIQSTAGH 179
      |||
Db 121 CRDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVMPAIQSSAQGH 180
QY 180 QLSALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNONRFCOLEIQRRLCLPRPCLA 239
      |||
Db 181 QLSALVTPASADGPCPNMSTAWGPCSTTCGLGIATRVSNONRFCOLEIQRRLCLSRPCLA 240
QY 240 ARSHSSWNSAF 250
      |||
Db 241 SRSHGSWNSAF 251
```

RESULT 4

```
US-10-112-267-78
; Sequence 78, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 78
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-267-78
```

```
Query Match 90.5%; Score 1303.5; DB 15; Length 250;
Best Local Similarity 90.0%; Pred. No. 1.8e-101;
Matches 225; Conservative 9; Mismatches 15; Indels 1; Gaps 1;
```

```
QY 2 RGSPLIHLATSFLCLLSMVCACLCTPCTCPMTTPPQCPQGVPLVLDGCGCKVCARRLG 61
      |||
Db 1 RGNPLIHLAISFLCILSMVYSQLCPAPCACPMTTPPQCPGVPLVLDGCGCCRCVARRLG 60
QY 62 ESCDHLHVCDBPSQGLVCQPGAGPGGHGAVCLLDEDDGSCSEVNGRRYLDGETFKPNCRVLC 121
      |||
Db 61 ESCDHLHVCDBPSQGLVCQPGAGPSGRGAVCLFEEDDGSCSEVNGRRYLDGETFKPNCRVLC 120
QY 122 RCDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDOGV-TPAIQSTAGHQ 180
      |||
Db 121 RCDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVMPAIQSSAQGHQ 180
QY 181 LSAVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNONRFCOLEIQRRLCLPRPCLA 240
      |||
Db 181 LSAVTPASADGPCPNMSTAWGPCSTTCGLGIATRVSNONRFCOLEIQRRLCLSRPCLA 240
QY 241 RSHSSWNSAF 250
      |||
Db 241 RSHGSWNSAF 250
```

RESULT 5

```
US-10-112-267-79
; Sequence 79, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 79
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-267-79
```

```
Query Match 90.2%; Score 1298.5; DB 15; Length 249;
Best Local Similarity 90.0%; Pred. No. 4.6e-101;
Matches 224; Conservative 9; Mismatches 15; Indels 1; Gaps 1;
```

```
QY 3 GSPLIHLATSFLCLLSMVCACLCTPCTCPMTTPPQCPQGVPLVLDGCGCKVCARRLGE 62
      |||
Db 1 GNPLIHLAISFLCILSMVYSQLCPAPCACPMTTPPQCPGVPLVLDGCGCCRCVARRLGE 60
QY 63 SCDHLHVCDBPSQGLVCQPGAGPGGHGAVCLLDEDDGSCSEVNGRRYLDGETFKPNCRVLCR 122
      |||
Db 61 SCDHLHVCDBPSQGLVCQPGAGPSGRGAVCLFEEDDGSCSEVNGRRYLDGETFKPNCRVLCR 120
QY 123 CDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDOGV-TPAIQSTAGHQ 181
      |||
Db 121 CDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVMPAIQSSAQGHQ 180
QY 182 SALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNONRFCOLEIQRRLCLPRPCLAAR 241
      |||
Db 181 SALVTPASADGPCPNMSTAWGPCSTTCGLGIATRVSNONRFCOLEIQRRLCLSRPCLASR 240
QY 242 SHSSWNSAF 250
      |||
Db 241 SHGSWNSAF 249
```

RESULT 6

```
US-10-112-267-80
; Sequence 80, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
```



```
Db      61 HLHVCDSQGLVQCQPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRLCRD 120
      :|||||
QY      126 GGFCLPLCSEDVRLPSWDCPRPRRIQVPGKCCPEWVCDQGV-TPAIQRSTAQGHQLSAL 184
      :|||||
Db      121 GGFCLPLCSEDVRLPSWDCPRPRRIQVPGRCCEWVCDQAVMQPAIQSSAQGHQLSAL 180
      :|||||
QY      185 VTPASADAPCPNMWSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLPRPCLARSHS 244
      :|||||
Db      181 VTPASADGPCPNWSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLSRPCLASRSHG 240
      :|||||
QY      245 SWNSAF 250
      :|||||
Db      241 SWNSAF 246
```

RESULT 9.

```
US-10-112-267-83
; Sequence 83, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 83
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-267-83
```

```
Query Match      88.9%; Score 1280.5; DB 15; Length 245;
Best Local Similarity 90.2%; Pred. No. 1.4e-99;
Matches 221; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY      7 HLHATSFLCLSMVCAQLCRPTCTCPWTPPQCPQGVPLVLDGCGCKVCARLGECDH 66
      :|||||
Db      1 HLHATISFLCILSMVSYQLCPADCACPWTPPQCPGVPLVLDGCGCCRVCARLGECDH 60
      :|||||
QY      67 LHVCDPSQGLVQCQPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRLCRD 126
      :|||||
Db      61 LHVCDPSQGLVQCQPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRLCRD 120
      :|||||
QY      127 GGFCLPLCSEDVRLPSWDCPRPRRIQVPGKCCPEWVCDQGV-TPAIQRSTAQGHQLSALV 185
      :|||||
Db      121 GGFCLPLCSEDVRLPSWDCPRPRRIQVPGRCCEWVCDQAVMQPAIQSSAQGHQLSALV 180
      :|||||
QY      186 TPAADAPCPNMWSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLPRPCLARSHS 245
      :|||||
Db      181 TPAADGPCPNWSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLSRPCLASRSHS 240
      :|||||
QY      246 WNSAF 250
      :|||||
Db      241 WNSAF 245
```

RESULT 10

```
US-10-112-267-84
; Sequence 84, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 84
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-267-84
```

```
Query Match      88.6%; Score 1276.5; DB 15; Length 244;
Best Local Similarity 90.2%; Pred. No. 3.1e-99;
Matches 220; Conservative 8; Mismatches 15; Indels 1; Gaps 1;
```

```
QY      8 HLHATSFLCLSMVCAQLCRPTCTCPWTPPQCPQGVPLVLDGCGCKVCARLGECDH 67
      :|||||
Db      1 HLHATISFLCILSMVSYQLCPADCACPWTPPQCPGVPLVLDGCGCCRVCARLGECDH 60
      :|||||
QY      68 HVCDBSQGLVQCQPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRLCRD 127
      :|||||
Db      61 HVCDBSQGLVQCQPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRLCRD 120
      :|||||
QY      128 FTCLPLCSEDVRLPSWDCPRPRRIQVPGKCCPEWVCDQGV-TPAIQRSTAQGHQLSALV 186
      :|||||
Db      121 FTCLPLCSEDVRLPSWDCPRPRRIQVPGRCCEWVCDQAVMQPAIQSSAQGHQLSALV 180
      :|||||
QY      187 PASADAPCPNMWSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLPRPCLARSHS 246
      :|||||
Db      181 PASADGPCPNWSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLSRPCLASRSHS 240
      :|||||
QY      247 NSAF 250
      :|||||
Db      241 NSAF 244
```

RESULT 11

```
US-10-112-267-85
; Sequence 85, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
```



```

: APPLICANT: Pennlca, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/10/112,267
: CURRENT FILING DATE: 2002-03-27
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
: PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 85
: LENGTH: 243
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-112-267-85

```

Query Match	88.1%;	Score 1268.5;	DB#15;	Length 243;
Best Local Similarity	90.1%;	Pred. No. 1.4e-98)		
Matches 219;	Conservative	8;	Mismatches 15;	Indels 1;
				Gaps 1;

```

qy      9 LATSFLCLSMVCAQLCRPTCTCPWTPPOCGVPVLJGCGCCKVCARRLGESCDHLH 68
        ||| ||||:|||| :||| | | ||||| |||| |||||:||||| |||||
db      1 LAISFLCILSMVYSQLCRAPCAPCAMPWPPOCGPVGVLJDGGCCRCVACARLGECDHLH 60

```

```
Qy      69 VCDPSQGLVCQPGAGPGGHGAVCLDEDDGSCSEVNGRRYLDGETFKPNCRYLCRCDGGF 128  
        |||||  
Db      61 VCDPSQGLVCQPGAGPGSGRGAVCLFEEDDGSCSEVNGRRYLDGETFKPNCRYLCRCDGGF 120
```

Oy	129 TCLPLCSEDVRLPSPWDCPRPKRIQYPGKCCPEWCVDG ³ -TPAIGRSTAQHQLSALVTP	187 : :
Db	121 TCLPLCSEDEVRLPSPWDCPPRPRIQYVGRCCEWCVDQA ⁴ MOPAIQPSSAQHQLSALVTP	180 : :

Qy	188	ASADAPCPNMSTAMGPCSTTCGLGATRVSNQNRFQOL	1	IORRLCLPRPCLARSHSSWN	247
	181	ASADGPCPNMSTAMGPCSTTCGLGATRVSNQNRFQOL <td>1 <td>IORRLCLSRPCLARSHGSWN <td>240</td> </td></td>	1 <td>IORRLCLSRPCLARSHGSWN <td>240</td> </td>	IORRLCLSRPCLARSHGSWN <td>240</td>	240

QY	248	SAF	250
	III		
DB	241	SAF	243

RESULT 12

; Sequence 86, Application US/10112267
; Publication No. US20030068678A1

APPLICANT: Botstein, David A.

APPLICANT: Goddard, Audrey

APPLICANT: Hillan, Kenneth J.
APPLICANT: Lawrence, David A.

APPLICANT: Pennica, Diane

APPLICANT: Wood, William I.

FILE REFERENCE: P1176R2

CURRENT FILING DATE: 2002-03-27

PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: US 60/063 704

;; PRIOR FILING DATE: 1997-10-29 1
;; PRIOR APPLICATION NUMBER: US 80/063,704
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:

PRIOR FILING DATE: 1998-02-04

```

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 86
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-267-86

```

Query Match	87.88;	Score 1264.5;	DB 15;	Length 242;
Best Local Similarity	90.18;	Pred. No. 3.1e-98;		
Matches 218; Conservative	8;	Mismatches 15;	Indels 1;	Gaps 1;

```
QY      10 LATSFLCLLSMCAOLCRHPTCTPWTBPQCPCPGVPLVLVDGCGCKVCARRLGESCDLHV 69
        |||||||:||||| || ||||| ||||| |||||:||||| ||||| |||||
Db       1 LAISFLCILSMVYSQLCPAPCACFPWTPPQCPCPGVPLVLVDGCGCCRCVARRLGESCDLHV 60
```

```

QY      70 CDP SQGLV CQPGAGPGG HGA VCLLEDDDGSC EVNGRRYLDGETFKPNC RYLCRCD DGGFT 129
        |||||
        ||||| : |||||
Db      61 CDP SQGLV CQPGAGP SG RGAVCLFEEDDGSC EVNGRRYLDGETFKPNC RYLCRCD DGGFT 120

```

QY 130 CLPLCEDVRLPSWDCPRPKRIQPCKCCPEWCDDGV-TPAIQRSTAQHLSALVTPA 188
||||| : ||||| | ||| : |||||
Dd 121 CULPCSEDRVLPSCDCCRPRRIQVPGRCCPEWVCDOAVMOPAIOPSSAQHLSALTPTA 180

QY	189	SADAPCPNMSTAWGPCSTTCGLGATRVSNQNRFCQLEIQRLCLPRPCLARSHSSWNS	248
Db	181	SADGPCPNMSTAWGPCSTTCGLGATRVSNQNRFCQLEIQRLCLSRPCLARSHSGSWNS	240

QY	249	AF	250
	11		
Db	241	AF	242

RESULT 13
US-10-112

US 10112267
; Sequence 87, Application US/10112267
; Publication No. US20030068678A1

GENERAL INFORMATION:

APPLICANT: Cohen, Robert
APPLICANT: Goddard, Andrew

;; APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.

APPLICANT: Levine, Arnold J.

APPLICANT: Roy, Margaret Ann

10 TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

CURRENT APPLICATION NUMBER: US/10/112,267
CURRENT FILING DATE: 2003-03-27

PRIOR FILING DATE: 1998-10-29

;; PRIOR APPLICATION NUMBER: 03 00/005,704
;; PRIOR FILING DATE: 1997-10-29

PRIOR FILING DATE: 1998-02-04

PRIOR FILING DATE: 1998-04-14

SEQ ID NO 87

; TYPE: PRT
: ORGANISM: H

18-197-7TT-0T-S0

Query Match	87.5%	Score 1260.5;	DB 15;	Length 241;
Best Local Similarity	90.0%;	Pred. No. 6.7e-98;		
Matches 217: Conservative	8;	Mismatches 15;	Indels 1;	Gaps 1;

11 11515CCELSMVCADCKRIFCI CPMIIPFQCFQSVFLEBDSGCCNVCARKLGGSCDHNLNC /0

Db	1	AISFICILSMVYSQLCPAPCACAPWTPEQCPGPVPLVLDDGGCCRCVACRLGESCDHLHVC	60
QY	71	DPSQGLVCQPGAGPGGHGAVCLLDEDDGSCSEVNGRRYLDGETFKPNCRYLCRCDDGGFTC	130
Db	61	DPSQGLVCQPGAGPSGRGAVCLFEEDDGSCSEVNGRRYLDGETFKPNCRYLCRCDDGGFTC	120
QY	131	LPICSEDEVRLPSWDCPRPKRIQVPGRCCPEWVCDQGV-TPAIQRSTAGHOLSALVTPAS	189
Db	121	LPICSEDEVRLPSWDCPRPRRIQVPGRCCPEWVCDQAVMQPAIQPSAQGHLSALVTPAS	180
QY	190	ADAPCPNMSTAMGPCSTTCGLGIATRVSNQNRFCQLEIQRLCLPRCLAARSHSSWNSA	249
Db	181	ADGPCPNMSTAMGPCSTTCGLGIATRVSNQNRFCQLEIQRLCLSRPCLASRSHGSWNSA	240
QY	250	F 250	
Db	241	F 241	

```

RESULT 14
US-10-112-267-89
; Sequence 89, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 89
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-267-89

```

Query Match	87.3%	Score 1257.5;	DB 15;	Length 239;
Best Local Similarity	90.4%	Pred. No. 1.2e-97;		
Matches 216;	Conservative 8;	Mismatches 14;	Indels 1;	Gaps 1;
QY	13	SFLCLLSMVCQAQLCRTPCTCPMTTPQCPQGVPLVLDGCGCKVCARRLGESCDHLHYCDP	72	
		: :		
Db	1	SFLCILSMVYSQLCPAFCACPMTPPQCPGPVPLVLDGCGCCRCARRLGESCDHLHYCDP	60	
QY	73	SQGLVYCQPGAGPGGHGAVCLLDEDDGSCSEVNGRRYLDGETFRKNCRYLCRCDDGGFTCLP	132	
Db	61	SQGLVYCQPGAGPGSGRGAVCLFEEDDGSCSEVNGRRYLDGETFRKNCRYLCRCDDGGFTCLP	120	
QY	133	LCSEDEVRLPSPWDCPRPKRIYVPGKCCPEWVCDQGV-TPAIDRSTAQGHQLSALVTPASAD	191	
Db	121	LCSEDEVRLPSPWDCPRPRRIYVPGRCCEPWVCDQAVMOPAIQPSAAGHQLSALVTPASAD	180	
QY	192	APCPNMSTAMGPCSTTCGLGIATRVSNONRRCOLEIQRRLCPRPCLARSHSSWNSAF	250	
Db	181	GPCPNMSTAMGPCSTTCGLGIATRVSNONRRCOLEIQRRLCPRPCLARSHSGSWNSAF	239	

```

RESULT 15
US-10-112-267-88
; Sequence 88, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Bolstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 88
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-267-88

```

	Query Match	87.3%;	Score 1257.5;	DB 15;	Length 240;
	.Best Local Similarity	90.4%;	Pred. No. 1.2e-97;		
	Matches 216;	Conservative	8;	Mismatches 14;	Indels 1; Gaps 1;
QY	13 SFLCLSMVCAQLCRTPCTCPWTPPOCGVPVLVLDDGCGCKVCARRLGESCDHLHYCDP	72			
Db	2 SFLCILSMVYSQLCPAPACACPWTPPQCPRPGVPLVLDDGCGCCRVCARLIGESCDHLHYCDP	61			
QY	73 SQGLVYCQPAGAPGGHGANVCLLDEDDGSCEVNGRRYLDETFFKPNCRVLCRCDDGGFTCLP	132			
Db	62 SQGLVYCQPAGAPSGRGANVCLFEEDDGSCEVNGRRYLDGETFFKPNCRVLCRCDDGGFTCLP	121			
QY	133 LCSEDEVRLPSWDCPRPKRIQVPGKCCEPEWVCDQGV-TPTAQRSTAQGHQLSALVTPASAD	191			
Db	122 LCSEDEVRLPSWDCPRPRRRIQVPGRCCEPEWVCDQAVMPEAIQPPSSAQGHQLSALVTPASAD	181			
QY	192 APCPNMWSTAWGPCSTTCGIGIATRVSNQNRFQOLEIQRLCLPRPCLAARSHSWSNSAF	250			
Db	182 GPCPNMWSTAWGPCSTTCGIGIATRVSNQNRFQOLEIQRLCLSRPCLASRSHGWSNSAF	240			

Search completed: July 25, 2003, 12:39:05
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:32:35 ; Search time 39 Seconds
(without alignments)
616.466 Million cell updates/sec

Title:	US-10-010-408-2
Perfect score:	1440
Sequence:	1 MRGSPLIHLATSPCLISM.....LCLPPTCLARSHSSWNSAF 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	564	39.2	349	2	A40551	connective tissue
2	550.5	38.2	348	2	A40578	beta IG-M2 protein
3	541.5	37.6	351	2	S20078	.NOV protein - chic
4	541	37.6	375	2	A41428	CEF-10 protein pre
5	519	36.0	357	2	I38069	gene novh protein
6	500.5	34.8	379	2	A35669	gene CYR61 protein
7	161	11.2	1620	2	T27283	hypothetical protee
8	146	10.1	1172	1	TSHUP2	thrombospondin 2 p
9	137.5	9.5	5376	2	T42215	zonadhesin - mouse
10	136.5	9.5	1036	2	T17405	scavenger receptor
11	136	9.4	1574	2	T13954	MEGF6 protein - ra
12	135	9.4	2321	2	S78549	notch3 protein - h
13	132.5	9.2	1111	2	T26972	hypothetical protee
14	131.5	9.1	593	1	GXHU	granulin precursor
15	131.5	9.1	1492	2	A40333	collagen alpha 1(''
16	131.5	9.1	2910	2	T42214	otogelin - mouse
17	130.5	9.1	1486	1	B40333	collagen alpha 1(I
18	130	9.0	1034	2	JC5598	mucin - rat
19	128.5	8.9	884	2	T18649	hypothetical prote
20	128.5	8.9	1496	1	CGHUV	collagen alpha 2(V
21	128	8.9	1178	1	A39804	thrombospondin pre
22	127	8.8	2531	2	A46019	notch-1 protein -
23	126.5	8.8	1466	1	CGHUTL	collagen alpha 1(I
24	126	8.8	1251	2	A57293	latent transformin
25	126	8.8	1964	2	T09059	notch4 - mouse
26	125.5	8.7	1487	1	CGHUC6	collagen alpha 1(I
27	124.5	8.6	810	2	T10756	Nel-homolog protei
28	124	8.6	837	2	A42112	mucin-like peptide
29	124	8.6	886	2	I50694	collagen alpha 1(I

30	124	8.6	1408	2	S16148	gene serrate prote
31	124	8.6	1487	2	B41182	collagen alpha 1(I
32	123	8.5	1444	2	T18856	angiogenesis inh
33	122.5	8.5	1188	2	D86236	protein F14N23.5
34	122.5	8.5	2201	2	A32160	tenascin-C - huma
35	122	8.5	304	2	A33274	insulin-like grow
36	122	8.5	305	2	I48601	insulin-like grow
37	122	8.5	305	2	JN0508	insulin-like grow
38	122	8.5	589	2	B38128	epithelin/granulin
39	121.5	8.4	282	2	S50031	prostaglandin-stimu
40	121.5	8.4	591	2	I48141	acroggranin - guine
41	121.5	8.4	2476	2	T34022	zonadhesin - pig
42	120	8.3	317	2	I46916	insulin-like grow
43	120	8.3	615	2	A05269	collagen alpha 1(I
44	120	8.3	1023	2	T30257	Igg Fc binding pro
45	120	8.3	2813	1	VWNU	von Willebrand fac

ALIGNMENTS

```

RESULT 1
A40551
connective tissue growth factor - human
C;Species: Homo sapiens (man)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
C;Accession: A40551; S44205
R;Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotendorst, G.R.
J. Cell Biol.,114, 1285-1294, 1991
A;Title: Connective tissue growth factor: a cysteine-rich mitogen secreted by human v
A;Reference number: A40551; MUID:91373462; PMID:1654338
A;Accession: A40551
A;Molecule type: mRNA
A;Residues: 1-349 <BRA>
A;Cross-references: GB:M92934; GB:M36965; GB:S56201; NID:g180923; PIDN:AAA91279.1; PI
R;Oemar, B.S.; Werner, A.; Yang, Z.; Garnier, J.M.; Gentz, R.; Luescher, T.F.
submitted to the EMBL Data Library, April 1994
A;Description: Differential cloning and expression of human connective tissue growth
A;Reference number: S44205
A;Accession: S44205
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-349 <OEM>
A;Cross-references: EMBL:X78947; NID:g474933; PID:g474934

Query Match          39.2%; Score 564; DB 2; Length 349;
Best Local Similarity 46.9%; Pred. No. 2.1e-35;
Matches 112; Conservative 28; Mismatches 81; Indels 18; Gaps 6;

QY      13 SFLLLSM---VCAQLCRTPCTCPWTP-PQCPQGVLPLDGCCKVCARRLGESCDHL 67
       |:|::: | | | | | | | | | | | | | | | | | | | | | | |
Db      12 AFVYLALCSRPAYGQNCSGPCPCDEPAAPRCAGVSLVDGCGCCRVCAKQLGELCTER 71

QY      68 HVCDSQGLVCQPGAGPGHGAVCLLDEDGSCSEVNGRRLDGFTFKPNCRVLCRDDGG 127
       ||:|| | | : | | | | | | | | | | | | | | | | | | | |
Db      72 DPDCPHKGLFCDFGSPANRKIGVCTA-KDGAPCIFGTVYRSGESFQSSCKYQCTCLDGA 130

QY      128 FTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCEPEWVCDQGVTPAIQRSTAOGHQLSAL--- 184
       |:| | | | | | | | | | | | | | | | | | | | | | | | |
Db      131 VGCMLPLCSMDVRLPSDPCCFFRRVKLPgkCEEWCDE--PKDQ--TIVGPALAAYRLE 185

QY      185 ----VTPASADAPCPNMSTANGBCSTTCGLGIATRVSNQNRFCQLEIQRRLLCPPCLA 239
       | | | | | | | | | | | | | | | | | | | | | | | | |
Db      186 DTFFPDPTMIIRANCLVQTTEWSACSKTCGMGISITRYVTNDNASCHLEKQSRLCMVRPCEA 244

RESULT 2
A40578
beta IG-M2 protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 01-Dec-2000
C;Accession: A40578; A53228
R;Brunner, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
```



```

A/Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)
C/Genetics:
A/Gene: GDB:THBS2; TSP2
A/Cross-references: GDB:128789; OMIM:188061
A/Map position: 6q27-6q27
C/Complex: homotrimer, disulfide linked
C/function:
A/Description: participates in cell migration and adhesion, and in platelet aggregation
C/Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; VC
C/Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F:1-18/Domain: signal sequence #status predicted <Sig>
F:19-1172/Product: thrombospondin 2 #status predicted <MAT>
F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F:553-588/Domain: EGF homology <EGF1>
F:652-691/Domain: EGF homology <EGF>
F:928-930/Region: cell attachment (R-G-D) motif
F:151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) #status predicted
F:167-226/Disulfide bonds: #status predicted
F:266,270/Disulfide bonds: Interchain #status predicted
F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match      10.1%; Score 146; DB 1; Length 1172;
Best Local Similarity 23.6%; Pred. No. 0.0016;
Matches 67; Conservative 22; Mismatches 85; Indels 110; Gaps 17;

QY 21 VCAQLCRTPCTCPWTPPQQPGVPLVLDDGCGCKVCARRLGESCDHLHVCDPSSQG----- 75
    :|: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 349 ICHQITCPATC-----ASPSVEG-ECCPSC-----LHSVDGEGWSPWA 388
                                     :|: | | | | | | | | | | | | | | | |

QY 76 -----LVCPGAGPGHGAVCLLDEDDGSCEVNGRRYLDETFFKPNCRVLCRD----- 124
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 389 EWTQCSVTGSGSTQGR-----SCDVTSMTCCLGPSIQTRACSLSKCDTRIRQ 436
                                     :|: | | | | | | | | | | | | | | | |

QY 125 DGGF-----TCLPLC--SEDEVRLPSMDCPRPKRIQVPCKCCPEWVCDOGVTPAIQRST 175
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 437 DGGWSHWSPWSSCSVTCGVGNITRIRLCNSPV---QMGGKNC----- 476
                                     :|: | | | | | | | | | | | | | | | |

QY 176 AQGHQLSALVTPASADAPCP---NMS--TAWGPCSTTCGYGT--ATRVSNQ-----NRF 222
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 477 ---KSGRETKACGACPIDGRWSPWSPWACTYTCAAGIGIRETRVCNSPEPOYGGA 532
                                     :|: | | | | | | | | | | | | | | | |

QY 223 COLEIQRLL-----CLPRPCLAARSHS-----SWNSAF 250
    | :|: | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 533 CVGDVQEROMCNKRSCPVDGCLSNPCPGAQCSSFPPDGSMSCGF 576
                                     :|: | | | | | | | | | | | | | | | |

RESULT 9
T42215
zonadhesin - mouse
N/Alternate names: sperm-specific membrane protein
C/Species: Mus musculus (house mouse)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T42215
R/Gao, Z.; Garbers, D.L.
J. Biol. Chem. 273, 3415-3421, 1998
A>Title: Species diversity in the structure of zonadhesin, a sperm-specific membrane pro
A/Reference number: 222080; MUID:98123114; PMID:9452463
A/Accession: T42215
A>Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-5376 <GAO>
A/Cross-references: EMBL:U97068; NID:g3327420; PID:g3327421; PIDN:AAC26680.1
C/Genetics:
A/Gene: Zan
A/Map position: 5
C/function:
A/Description: functions in multiple cell adhesion processes
A/Note: found exclusively on the apical region of the sperm head
C/Keywords: cell adhesion

```

[illegible]

```

RESULT 10
T17405
scavenger receptor cysteine-rich protein precursor - sea urchin (Strongylocentrotus p
C/Species: Strongylocentrotus purpuratus (purple urchin)
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C/Accession: T17405
R/Pancer, Z.; Rast, J.P.; Davidson, E.H.
Immunogenetics 49, 773-786, 1999
A/Title: Origins of immunity: transcription factors and homologs of effector genes of
A/Reference number: Z18253; MUID:99328904; PMID:10388804
A/Accession: T17405
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Residues: 1-1036 <PAN>
A/Cross-references: EMBL:AF076513; NID:g4164530; PID:g4164531; PIDN:AAD05493.1

Query Match          9.5%; Score 136.5; DB 2; Length 1036;
Best Local Similarity 26.4%; Pred. No. 0.0075;
Matches 63; Conservative 15; Mismatches 72; Indels 89; Gaps 15;

QY 21 VCAQLCRTPCTCPWTPPQCPCQGVPLVLDGCGCKVCARRLGESCDHLHVCDP---SQGLV 77
   :| ||| | |||: |||: ||| | ||| | |||: |||:
DB 54 ICPLLCRAGCF-----CPEGLVKDRDG-----GDRCIHLDDQCQDRHCPEGMT 95

QY 78 --QCPGAGPGHGA VCLLDEDDGSCVNGRRYLDGETFKPNC RVLGRC-----DDGG 127
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 96 FDECGSGCGP-----ASCE-----NLPRDIPRICRAGCFCEGLVKDQDDGG 137

QY 128 FTCLPL-----CSEDEV-----RLPSWDCPRPKRIQVPKGCPE-WVCD 164
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 138 DRCIPLDHCHQDRHCPDGMAYDECGSGCGPFSCDNLPSYICPRICR---AGCFCEGLVKD 194

QY 165 Q-GVTPAIRQSTAGQHQLSALVTPASADAPCPNWSSTAWGPCSTTCGLGIATRVSNQNR 222
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 195 QDGGDRCIPLDQCQGDPLNCPVT-----GQWFVDW-DASYEC---LSQLSQSNPF 240

RESULT 11
T13954
MEGF6 protein - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C/Accession: T13954
R/Nakajima, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A/Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A/Reference number: Z14126; MUID:98360089; PMID:9693030
A/Accession: T13954
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Residues: 1-1574 <NAK>
A/Cross-references: EMBL:AB011532; NID:g3449293; PIDN:BAA32462.1; PID:g3449294
A/Experimental source: strain Sprague-Dawley; brain
C/Genetics:
A/Gene: MEGF6

```

Query Match 9.4%; Score 136; DB 2; Length 1574;
Best Local Similarity 24.1%; Pred. No. 0.011;
Matches 79; Conservative 18; Mismatches 117; Indels 114; Gaps 21;

QY 1 MRGSPLHLATSLFLCL-----LSMVAQLCRT-----PCNCP--WT----- 35
Db 503 LRGE---HTLLEKFCVLDHSGHDCSLTCD-GRNGTCTFPGDGCDCEGWTGIICNET 558
QY 36 -PP-----QCPQGV--PLVLDC-----GCCKVCARR 59
Db 559 CPPDTFGKNSSPCTCQNGGTCDPVLGACRCRPGVSGAHCEDGCKPRGFYKHKCRKKCHCA 618
QY 60 LGESCDHLH---VCDPSQ-----GLVQPGA-GPG-----GHGAVCLDEDDGSC 101
Db 619 NRGCRHLRYGACLCDPGLYGRFCHLACPPWAFGPGCSEDCLEQSHTRSC--NPKDGSCS 676
QY 102 ----VNGRR---YLDGETFKPNCRVLCRCDDGGFTCLPLCSEDEVRLPSWDCPRPKRIQVP 154
Db 677 CKAGFQGERCOAECESGFFGPGCRHRTCP-GVACDPVSGE---CRTQCPPGYQGEDC 731
QY 155 GKCCPEWVCDQGVTPAIQRSTAQHQLS--ALVTP-----ASADAPCPNWSTAWGPCST 206
Db 732 GQECPPVGTFFGVNCSGSCSVGAPCHRTGECICPPGKTGEDCGADCEGRWGLGCOEICP 791
QY 207 TCGLGIAFRVSNQNRFCOLEIQRRLCP 234
Db 792 ACEHGAS-----CNPETGTCLCLP 810

RESULT 12

S78549

notch3 protein - human

C:Species: Homo sapiens (man)

C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 08-Sep-2002

C:Accession: S78549; S71825

R:Joutel, A.; Tournier-Lasserre, E.

submitted to the EMBL Data Library, April 1997

A:Reference number: S78549

A:Accession: S78549

A:Molecule type: mRNA

A:Residues: 1-2321 <J001>

A:Cross-references: EMBL:U97669; NID:g2668591; PIDN:AAB91371.1; PID:g2668592

R:Joutel, A.; Corpechot, C.; Ducros, A.; Vahedi, K.; Chabrier, H.; Mouton, P.; Alamowicz

x, M.M.; Weissenbach, J.; Bach, J.F.; Bousser, M.G.; Tournier-Lasserre, E.

Nature 383, 707-710, 1996

A:Title: Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke

A:Reference number: S71825; MUID:97032728; PMID:8878478

A:Accession: S71825

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 67-113;138-194;268-333,'G',335-346;536-613;716-765;1240-1279;1815-1888 <J002

A:Cross-references: EMBL:U97669

C:Genetics:

A:Gene: notch3

A:Map position: 19p13.1

C:Function:

A:Description: may be involved in pathogenesis of CADASIL, causing a type of stroke and

C:Superfamily: notch protein; ankyrin repeat homology; EGF homology

C:Keywords: tandem repeat; transmembrane protein

F:123-155/Domain: EGF homology <EGX1>

F:162-194/Domain: EGF homology <EGF1>

F:240-271/Domain: EGF homology <EGX2>

F:318-349/Domain: EGF homology <EGF>

F:473-504/Domain: EGF homology <EGX3>

F:853-884/Domain: EGF homology <EGF3>

F:928-959/Domain: EGF homology <EGX4>

F:1838-1870/Domain: ankyrin repeat homology <AN1>

F:1871-1903/Domain: ankyrin repeat homology <AN2>

F:1905-1937/Domain: ankyrin repeat homology <AN3>

F:1938-1970/Domain: ankyrin repeat homology <AN4>

F:1971-2003/Domain: ankyrin repeat homology <AN5>

Query Match

9.4%; Score 135; DB 2; Length 2321;

Best Local Similarity 22.9%; Pred. No. 0.018;
Matches 57; Conservative 13; Mismatches 67; Indels 112; Gaps 12;

QY 30 CTCF--WTPPQC-----PQGVPLVLDGCGCCKVCARRLGESCDH-LHVCDP 72
Db 796 CSCPQGWQGRCCQDDVDECAGPAPCPGHGICITNLAGSFSCCTCHGTYTGPSCDDINDCDP 855
QY 73 S---QGLVQPGAAG-----PGHGAVCLDEDD-----GSCEVNGRRYLDGETFK 114
Db 856 NPCLNGSCQDGVGSFSCSCLPFGAGPCARDVDECLSNPCPGTCTDHVASF----- 908
QY 115 PNCRVLCRCD--GGFTC---LPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGVTP 169
Db 909 ----TCTCPPGYGGFHCEQDLPCDS-----PSSCFNGCTCVDGV-- 943
QY 170 AIQRSTAQGHQLSALVTPASADAPCPMNSTAWGPCSTTCGLGIAFRVSNQNRFCOLEIQR 229
Db 944 -----NSFSLCRPGYTGHAHQH-----EA 963
QY 230 RLCLPRPCL 238
Db 964 DPCLSRPCL 972

RESULT 13

T26972

hypothetical protein Y47H9C.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T26972

R:Harris, B.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20293

A:Accession: T26972

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1111 <WIL>

A:Cross-references: EMBL:AL032657; PIDN:CAA21739.1; GSPDB:GN00019; CESP:Y47H9C.4

A:Experimental source: clone Y47H9C

C:Genetics:

A:Gene: CESP:Y47H9C.4

A:Map position: 1

A:Introns: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol

Query Match 9.2%; Score 132.5; DB 2; Length 1111;
Best Local Similarity 26.1%; Pred. No. 0.016;
Matches 61; Conservative 19; Mismatches 93; Indels 61; Gaps 15;

QY 22 CAQLCRT-----PCTCP--WTPPQCQGVPLVLDGCGCCVCARRLGESCDHL 67
Db 470 CAHQCCQNRGVGCDGADGKCQCDRGMTCGRCEHNCRPADTFGANCERKCKCPKIGCDPI 529
QY 68 H---VCDPS-QGLVQPGAAGPGGHGAVCLDEDDGSC- VNGRRYLDGETFKPNCRVLCR 122
Db 530 TGECTCPAGLOGANCNDIGCPESYGPCKL-----HCKCVNGK--CDKETGE-----CT 576
QY 123 CDDG--GFTCLPLCSE-----DVRLPAMD-----CP-RPKRIQVPGKCP- 160
Db 577 CQPGFFGSDCSTCSKKGYESCELSCPDASCSKQTKGKLCPLGTGKGVSCDQKDPNT 636
QY 161 --WVCDQGVTPAIQRSTAQGHQLSALVTPASA---DAPCPMNSTAWGPCSTTC 208
Db 637 FGFLCQETVTPSPCASTDPKNGVCLSCPSSGIIHCEHNCAGSYGDG-CQQVC 689

RESULT 14

GYHU

granulin precursor [validated] - human

N:Alternate names: epithelin

N:Contains: granulin A; granulin B; granulin C; granulin D; granulin E; granulin F; g

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1992 #sequence_revision 03-May-1996 #text_change 08-Dec-2000

C:Accession: JCI1284; A38128; A38118; A36698; B36698; C36698; D36698; A56873
R:Bhandari, V.; Bateman, A.
Biochem. Biophys. Res. Commun. 188, 57-63, 1992
A:Title: Structure and chromosomal location of the human granulin gene.
A:Reference number: JCI1284; MUID:93038704; PMID:1417868
A:Accession: JCI1284
A:Molecule type: DNA
A:Residues: 1-593 <BHA>
R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro, G.J.
J. Biol. Chem. 267, 13073-13078, 1992
A:Title: The epithelin precursor encodes two proteins with opposing activities on epithelial cells.
A:Reference number: A38128; MUID:92317004; PMID:1618805
A:Accession: A38128
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-593 <PLO>
A:Cross-references: GB:M62320; NID:g31192; PIDN:CAA44196.1; PID:g31193
R:Bhandari, V.; Palfree, R.G.E.; Bateman, A.
Proc. Natl. Acad. Sci. U.S.A. 89, 1715-1719, 1992
A:Title: Isolation and sequence of the granulin precursor cDNA from human bone marrow.
A:Reference number: A38118; MUID:9219253; PMID:1542665
A:Accession: A38118
A:Molecule type: mRNA
A:Residues: 1-406, 'R', 408-433, 'G', 435-453, 'G', 455-459, 'Q', 461-546, 'A', 548-566, 'R', 568-593
A:Cross-references: GB:M75161; NID:g183612; PIDN:AAA58617.1; PID:g183613
A:Note: this sequence has been revised in reference JCI1284
R:Bateman, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.
Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990
A:Title: Granulins, a novel class of peptide from leukocytes.
A:Reference number: A36698; MUID:91097544; PMID:2268320
A:Accession: A36698
A:Molecule type: protein
A:Residues: 281-336 <BAT>
A:Note: this protein was purified and characterized as granulin A
A:Accession: B36698
A:Molecule type: protein
A:Residues: 206-218, 'H', 220-233 <BA2>
A:Note: this protein was purified and characterized as granulin B
A:Accession: C36698
A:Molecule type: protein
A:Residues: 364-367, 'X', 369-385, 'H', 387-396 <BA3>
A:Note: this protein was purified and characterized as granulin C
A:Accession: D36698
A:Molecule type: protein
A:Residues: 442-446, 'XDTS', 456-458, 'DG' <BA4>
R:Kardana, A.; Bagshawe, K.D.; Coles, B.; Read, D.; Taylor, M.
Br. J. Cancer 67, 666-692, 1993
A:Title: Characterisation of UGP and its relationship with beta-core fragment.
A:Reference number: A56873; MUID:93229246; PMID:8471426
A:Accession: A56873
A:Molecule type: protein
A:Residues: 281-283, 'X', 285-289, 'S', 291-295 <KAR>
A:Experimental source: urine
A:Note: sequence extracted from NCBI backbone (NCBIP:129524)
C:Genetics:
A:Gene: GDB:GRN
A:Cross-references: GDB:136006; OMIM:138945
A:Map position: 17pter-17qter
A:Introns: 46/3; 88/3; 117/1; 154/3; 200/1; 236/3; 279/1; 311/3; 393/3; 471/3; 548/3
C:Superfamily: granulin
C:Keywords: glycoprotein; tandem repeat
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-593/Product: granulin #status predicted <MAT>
F:18-593/Product: progranulin #status predicted <PRO>
F:18-44/Product: paraganulin #status experimental <PGR>
F:58-113/Product: granulin G #status predicted <GRG>
F:123-179/Product: granulin F #status predicted <GRF>
F:206-261/Product: granulin B #status experimental <GRB>
F:281-336/Product: granulin A #status experimental <GRA>
F:364-417/Product: granulin C #status experimental <GRC>
F:442-496/Product: granulin D #status predicted <GRD>
F:518-573/Product: granulin E #status predicted <GRE>
F:368/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 9.1%; Score 131.5; DB 1; Length 593;
Best Local Similarity 24.6%; Pred. No. 0.011;
Matches 62; Conservative 15; Mismatches 78; Indels 97; Gaps 16;
QY 32 CPWTPQCPQGVPLVLDGCGCKV-----CARLGESC--DHLHVCDPSPGLVCQPGAG 83
DB 208 CPDARSRCP-----DGSTCCELPSRGKYGCCPMPNATCCSDHLHCC--PQDTVCDL--- 255
QY 84 PGGHGAVCLDDED-----DSCSEVNGRRYLDGETFRPNCRVLCRCDGGE 128
DB 256 ---IQSKCLSKENATTDLLTKLPAHYVDVVC-----DMEVSCPDPGYTCRCRLQSGAW 304
QY 129 TCLP---LCSEVDVRLPSWDCPRPKIQVEGKCCP-----EWVCDQGVTPAIQRSTA 176
DB 305 GCCPFTQAVCCED-----HIH-----CCPAGFTCDTQKGTCEQGPHQVPMMEKA 348
QY 177 QGHQLSALVTPAS--ADAPCPNWSFRWGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLP 234
DB 349 PAH--LSLPDPQALKRDVPCDNVSSC--PSSDTC-----CQLTSGEWGCCP 390
QY 235 RP---CLARSH 243
DB 391 IPEAVCCSDHQH 402
RESULT 15
A40333
C:Species: Xenopus laevis (African clawed frog)
C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Jul-1999
C:Accession: A40333
R:Su, M.W.; Suzuki, H.R.; Bleker, J.J.; Solursh, M.; Ramirez, F.
J. Cell Biol. 115, 565-575, 1991
A:Title: Expression of two nonallelic type II procollagen genes during Xenopus laevis
A:Reference number: A40333; MUID:92011898; PMID:1918153
A:Accession: A40333
A:Molecule type: nucleic acid sequence not shown
A:Residues: 1-1492 <SUA>
A:Cross-references: GB:M63596
A:Note: this sequence is presented as substitutions relative to another sequence in a
es they replace; the appropriate interpretation of the sequence figure was reconstruc
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F:37-96/Domain: von Willebrand factor type C repeat homology <VWC>
F:1263-1492/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
Query Match 9.1%; Score 131.5; DB 2; Length 1492;
Best Local Similarity 36.7%; Pred. No. 0.024;
Matches 33; Conservative 12; Mismatches 26; Indels 19; Gaps 5;
QY 94 DEDD---GSCEVNGRRYLDGETFRPNCRVLCRCDGGEFTCLPCESEDVRLPSWDCPRPK 149
DB 28 DEEDVLDTGSCVQDQGRYSKDKVWKPEPCQICVCDTG---TVLCDEIIICEESKDCP--- 80
QY 150 RIQVP-GKCCPEWVCDQGVTPAIQRSTAOG 178
DB 81 NAEIPFGECCP-----ICPTQSGSTSSG 103

Search completed: July 25, 2003, 12:38:00
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:25:50 ; Search time 23 Seconds

(without alignments)
511.160 Million cell updates/sec

Title: US-10-010-408-2

Perfect score: 1440

Sequence: 1 MRGSPLIHLATSFLLCLSM.....LCLPRPCLARSHSSWNSAF 250

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1390	96.5	250	1 CTGL_RAT	Q9jhc6 rattus norv
2	1308.5	90.9	251	1 CTGL_MOUSE	Q9z0g4 mus musculu
3	1064	73.9	250	1 CTGL_HUMAN	O76076 homo sapien
4	568	39.4	349	1 CTGF_BOVIN	O18739 bos taurus
5	564	39.2	349	1 CTGF_HUMAN	P29279 homo sapien
6	550.5	38.2	347	1 CTGF_RAT	Q9r1e9 rattus norv
7	550.5	38.2	348	1 CTGF_MOUSE	P29268 mus musculu
8	541.5	37.6	351	1 NOV_CHICK	P28686 gallus gall
9	541	37.6	375	1 CE10_CHICK	P19336 gallus gall
10	533	37.0	349	1 CTGF_PIG	O19113 sus scrofa
11	532.5	37.0	353	1 NOV_COTUA	P42642 coturnix co
12	529.5	36.8	354	1 NOV_MOUSE	O64299 mus musculu
13	519	36.0	357	1 NOV_HUMAN	P48745 homo sapien
14	516	35.8	351	1 NOV_RAT	Q9qzq5 rattus norv
15	512	35.6	343	1 NOV_XENLA	P1609 xenopus lae
16	504	35.0	379	1 CYR6_RAT	Q9es72 rattus norv
17	500.5	34.8	379	1 CYR6_MOUSE	P18406 mus musculu
18	498.5	34.6	381	1 CYR6_HUMAN	O00622 homo sapien
19	146	10.1	1172	1 TSP2_HUMAN	P35442 homo sapien
20	142	9.9	2282	1 ZAN_RABIT	P57999 oryctolagus
21	139	9.7	184	1 ESM1_HUMAN	Q9ng30 homo sapien
22	138	9.6	810	1 NEL1_HUMAN	Q92832 homo sapien
23	137.5	9.5	5376	1 ZAN_MOUSE	O88799 mus musculu
24	135	9.4	2321	1 NTC3_HUMAN	Q9um47 homo sapien
25	133	9.2	830	1 SREC_HUMAN	Q14162 homo sapien
26	131.5	9.1	593	1 GRN_HUMAN	P28799 h granulin
27	130	9.0	867	1 SSPO_BOVIN	P98167 bos taurus
28	130	9.0	2482	1 VWF_PIG	Q28833 sus scrofa
29	128.5	8.9	281	1 IBP7_MOUSE	O61581 mus musculu
30	128.5	8.9	1459	1 CA12_MOUSE	P28481 mus musculu
31	128.5	8.9	1496	1 CA25_HUMAN	P05997 homo sapien
32	128	8.9	1178	1 TSP2_CHICK	P35440 gallus gall
33	127.5	8.9	2812	1 ZAN_HUMAN	Q9y493 homo sapien

34	127	8.8	2531	1 NTC1_MOUSE	Q01705 mus musculu
35	126.5	8.8	1466	1 CA13_HUMAN	P02461 homo sapien
36	126	8.8	1964	1 NTC4_MOUSE	P31695 mus musculu
37	125.5	8.7	588	1 GRN_RAT	P23785 r granulin
38	125	8.7	1629	1 ATG9_HUMAN	Q9P2n4 homo sapien
39	124.5	8.6	810	1 NEL1_RAT	Q62919 rattus norv
40	124	8.6	837	1 MUC1_RAT	P98089 rattus norv
41	124	8.6	1262	1 CA13_CHICK	P12105 gallus gall
42	124	8.6	1408	1 SERR_DROME	P18168 drosophila
43	123	8.5	2319	1 NTC3_RAT	Q9r172 rattus norv
44	122.5	8.5	2201	1 TENA_HUMAN	P24821 homo sapien
45	122.5	8.5	2911	1 FBN2_HUMAN	P35556 homo sapien

ALIGNMENTS

RESULT 1					
CTGL_RAT					
ID	CTGL_RAT	STANDARD;	PRT;	250 AA.	
AC	Q9JHC6;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Connective tissue growth factor-like protein precursor (CTGF-L) (WNT1				
DE	Inducible signaling pathway protein 2) (WISP-2) (CCN family protein				
DE	COP-1).				
GN	WISP2 OR CTGFL OR COP1.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98414629; PubMed=9742130;				
RA	Zhang R., Averbouch L., Zhu W., Zhang H., Jo H., Dempsey P.J.,				
RA	Coffey R.J., Pardee A.B., Liang P.;				
RT	"Identification of rCop-1, a new member of the CCN protein family, as				
RT	a negative regulator for cell transformation.";				
RL	Mol. Cell. Biol. 18:6131-6141(1998).				
CC	-I- FUNCTION: May play an important role in modulating bone turnover.				
CC	Promotes the adhesion of osteoblast cells and inhibits the binding				
CC	of fibrinogen to integrin receptors. In addition, inhibits				
CC	osteocalcin production (By similarity).				
CC	-I- SUBCELLULAR LOCATION: Secreted (Probable).				
CC	-I- SIMILARITY: Contains 1 IGFBP domain.				
CC	-I- SIMILARITY: Contains 1 VWFC domain.				
CC	-I- SIMILARITY: Contains 1 TSP type-1 domain.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; AF259981; AAF69011.1; -				
DR	InterPro; IPR000867; Insl_gro_fac_pr.				
DR	InterPro; IPR000884; TSP1.				
DR	InterPro; IPR001007; VWF_C.				
DR	Pfam; PF00219; IGFBP; 1.				
DR	Pfam; PF00090; tsp_1; 1.				
DR	Pfam; PF00093; vwc; 1.				
DR	SMART; SM00121; IB; 1.				
DR	SMART; SM00209; TSP1; 1.				
DR	SMART; SM00214; VWC; 1.				
DR	PROSITE; PS00222; IGF_BINDING; 1.				
DR	PROSITE; PS50092; TSP1; 1.				
DR	PROSITE; PS01208; VWF_C_1; 1.				
DR	PROSITE; PS50184; VWF_C_2; 1.				
KW	Signal.				
FT	SIGNAL 1 23 POTENTIAL.				

```

FT CHAIN 24 250 CONNECTIVE TISSUE GROWTH FACTOR-LIKE
FT DOMAIN 26 96 PROTEIN.
FT DOMAIN 98 164 IGFBP.
FT DOMAIN 194 238 VMFC.
FT CARBOHYD 196 196 TSP TYPE-1.
SO SEQUENCE 250 AA; 27005 MW; 9A147074626BCA47 CRC64; N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match 96.5%; Score 1390; DB 1; Length 250;
Best Local Similarity 96.8%; Pred. No. 2.3e-98;
Matches 242; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 MRGSPILHLATSFLLSLMWCAQLCRPTCTPWTPOGPGVPLVLDGCGCKVCARRL 60
Db 1 MRGSPILRLATSFLLSLMWCAQLCRPTCTPWTPOGPGVPLVLDGCGCKVCARRL 60
OY 61 GESCDHLHVCDPSPGGLVCPGAGPGHGAVALDEDDGSCENVNGRRYLDGETFKPNCRVL 120
Db 61 TESCEHLHVCPSGGLVCPGAGPGHGAVALDEDDGSCENVNGRRYLDGETFKPNCRVL 120
OY 121 CRDDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVTPAIQSTAGQH 180
Db 121 CRDDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVTPAIQSTAGQH 180
OY 181 LSALVTPASADAPCPNWTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLPRCLAA 240
Db 181 LSALVTPASADAPCPNWTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLPRCLAA 240
OY 241 RSHSSWNSAF 250
Db 241 RSHSSWNSAF 250

RESULT 2
CTGL_MOUSE
ID CTGL_MOUSE STANDARD; PRT; 251 AA.
AC 0920G4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Connective tissue growth factor-like protein precursor (CTGF-L)
GN (WNT1 inducible signaling pathway protein 2) (WISP-2).
OS WISP2 OR CTGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=99061933; PubMed=9843955;
RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
RA Gurney A.L., Botstein D., Levine A.J.;
RT "WISP genes are members of the connective tissue growth factor family
RT that are up-regulated in wnt-1-transformed cells and aberrantly
RT expressed in human colon tumors."
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=99287915; PubMed=10358067;
RA Kumar S., Hand A.T., Connor J.R., Dodds R.A., Ryan P.J., Trill J.J.,
RA Fisher S.M., Nuttall M.E., Lipshutz D.B., Zou C., Hwang S.M.,
RA Voita B.J., James I.E., Rleman D.J., Gowen M., Lee J.C.;
RT "Identification and cloning of a connective tissue growth factor-like
RT cDNA from human osteoblasts encoding a novel regulator of osteoblast
RT functions."
RL J. Biol. Chem. 274:17123-17131(1999).
CC -1- FUNCTION: May play an important role in modulating bone turnover.
CC Promotes the adhesion of osteoblast cells and inhibits the binding
CC of fibrinogen to integrin receptors. In addition, inhibits
```

```

CC osteocalcin production (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 VMFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF100778; AAC96320.1; -.
CC EMBL; AF126063; AAD18058.1; -.
CC MGD; MGI:1328326; Wisp2.
CC InterPro; IPR000867; Insl_gro_fac_pr.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR001007; VMF_C.
CC Pfam; PF00219; IGFBP; 1.
CC Pfam; PF00090; tsp_1; 1.
CC Pfam; PF00093; vmc; 1.
CC SMART; SM00121; IB; 1.
CC SMART; SM00209; TSP1; 1.
CC SMART; SM00214; VMC; 1.
CC PROSITE; PS00222; IGF_BINDING; 1.
CC PROSITE; PS50092; TSP1; 1.
CC PROSITE; PS01208; VMFC_1; 1.
CC PROSITE; PS50184; VMFC_2; 1.
CC Signal.
CC FT SIGNAL 1 23 POTENTIAL.
CC FT CHAIN 24 251 CONNECTIVE TISSUE GROWTH FACTOR-LIKE
CC FT DOMAIN 26 96 IGFBP.
CC FT DOMAIN 98 164 VMFC.
CC FT CARBOHYD 196 197 TSP TYPE-1.
CC FT SEQUENCE 251 AA; 27095 MW; 893E163F6E5C7FC CRC64; N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match 90.9%; Score 1308.5; DB 1; Length 251;
Best Local Similarity 90.0%; Pred. No. 3.2e-92;
Matches 226; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

OY 1 MRGSPILHLATSFLLSLMWCAQLCRPTCTPWTPOGPGVPLVLDGCGCKVCARRL 60
Db 1 MRGNPLIHLATSFLLSLMWVYSQLCPAPACPMWTPPOGPGVPLVLDGCGCKVCARRL 60
OY 61 GESCDHLHVCDPSPGGLVCPGAGPGHGAVALDEDDGSCENVNGRRYLDGETFKPNCRVL 120
Db 61 GESCDHLHVCDPSPGGLVCPGAGPGHGAVALDEDDGSCENVNGRRYLDGETFKPNCRVL 120
OY 121 CRDDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVTPAIQSTAGQH 179
Db 121 CRDDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVTPAIQSTAGQH 180
OY 180 QLSALVTPASADAPCPNWTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLPRCLAA 239
Db 181 QLSALVTPASADAPCPNWTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRRLCLSRPCLAA 240
OY 240 ARSHSSWNSAF 250
Db 241 SRSHGWSNSAF 251

RESULT 3
CTGL_HUMAN
ID CTGL_HUMAN STANDARD; PRT; 250 AA.
AC 076076;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Connective tissue growth factor-like protein precursor (CTGF-L) (WNT1
```

DE Inducible signaling pathway protein 2) (WISP-2) (Connective tissue
DE growth factor-related protein 58).
GN WISP2 OR CTGFL OR CT58 OR CCN5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99287915; PubMed=10358067;
RA Kumar S., Hand A.T., Connor J.R., Dodds R.A., Ryan P.J., Trill J.J.,
RA Fisher S.M., Nuttall M.E., Lipshutz D.B., Zou C., Hwang S.M.,
RA Volta B.J., James I.E., Rieman D.J., Gowen M., Lee J.C.;
RT "Identification and cloning of a connective tissue growth factor-like
RT cDNA from human osteoblasts encoding a novel regulator of osteoblast
RT functions.";
RL J. Biol. Chem. 274:17123-17131(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Rowles J., Gendler S.;
RT "CT58, a new member of the connective tissue growth factor family,
RT interacts with the breast cancer associated mucin MUC1.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=99061933; PubMed=9843955;
RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
RA Gurney A.L., Botstein D., Levine A.J.;
RT "WISP genes are members of the connective tissue growth factor family
RT that are up-regulated in wt-1-transformed cells and aberrantly
RT expressed in human colon tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehaeslaio M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Keltman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -i- FUNCTION: May play an important role in modulating bone turnover.
CC Promotes the adhesion of osteoblast cells and inhibits the binding
CC of fibrinogen to integrin receptors. In addition, inhibits
CC osteocalcin production.
CC -i- SUBCELLULAR LOCATION: Secreted (Probable).
CC -i- TISSUE SPECIFICITY: Expressed in primary osteoblasts, fibroblasts,
CC ovary, testes, and heart.
CC -i- SIMILARITY: Contains 1 IGFBP domain.
CC -i- SIMILARITY: Contains 1 VWFC domain.
CC -i- SIMILARITY: Contains 1 TSP type-1 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF083500; AAC70350.1; -;
DR EMBL; AF074604; AAC26794.1; -;
DR EMBL; AF100780; AAC96322.1; -;
DR EMBL; AL139352; CAB94788.1; -;
DR EMBL; BC017782; AAH17782.1; -;
DR Genew; HGNC:12770; WISP2.
DR MIM; 603399; -;
DR GO; GO:0005625; C:soluble fraction; TAS.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0007048; P:oncogenesis; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 250 CONNECTIVE TISSUE GROWTH FACTOR-LIKE
FT PROTEIN.
FT DOMAIN 26 96 IGFBP.
FT DOMAIN 98 164 VWFC.
FT DOMAIN 194 238 TSP TYPE-1.
SQ SEQUENCE 250 AA; 26825 MW; CA99837EF42FEEAC CRC64;

Query Match 73.9%; Score 1064; DB 1; Length 250;
Best Local Similarity 73.6%; Pred. No. 9.2e-74;
Matches 184; Conservative 17; Mismatches 49; Indels 0; Gaps 0;

QY 1 MRGSPILHLATSFLLCLSMVCAQICRTPTCTCPWTPPQCQGVPLVLDGCGCKVCARRL 60


```
DB 1 MRGTPKTHLAFSLCLLSKVRTOQLCTPTCTCPWPPRCPGLVPLVDGCGCCRCVARRL 60
OY 61 GESCDHLVCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCSEVNGRRYLDGETFKPNCRVL 120
DB 61 GEPDQHLVCDASQGLVCPGAGPGGRGALCLLAEDDSSCEVNGRRYREGETFQPHCSIR 120
OY 121 CRCDGCGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVTPAIQRTAQGHQ 180
DB 121 CRCEDDGCGFTCVPLCSESDVRLPSWDCPHPRVEVLGKCCPEWVCDOGVTPAIQRTAQGHQ 180
OY 181 LSLVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRLCLPRPCLAA 240
DB 181 FSGLVSSLPPGVPCEPWSTAWGPCSTTCGLGMATRVSNQNRFCRLQRLCLSRPCPPS 240
OY 241 RSHSSWNSAF 250
DB 241 RGRSPNSAF 250

RESULT 4
CTGF_BOVIN STANDARD; PRT; 349 AA.
ID CTGF_BOVIN STANDARD; PRT; 349 AA.
AC 018739; Q9GL71;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OX Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Aorta;
RA Lillensiek B., Lin Z., Fotsis T., Schlanski M., Blierhaus A.,
RA Kanitz M., Kaufmann G., Schweigener L., Ziegler R., Nawroth P.P.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Liver;
RA Mathahs M., Schwitters C., Hove M., Rupp S., Erondu N.E.;
RT "Bovine connective tissue growth factor, organization of the
RT chromosomal gene and demonstration of promoter activity.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGF1P domain.
CC -1- SIMILARITY: Contains 1 VMFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF000137; AAB65596.1; -
DR EMBL; AF309555; AAC30290.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac_pr.
```

```
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VMFC_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGF1P; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vmc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VMFC_1; 1.
DR PROSITE; PS50184; VMFC_2; 1.
DR Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
KW SIGNAL.
FT 1 26
FT CHAIN 27 349
FT DOMAIN 33 100
FT DOMAIN 101 167
FT DOMAIN 198 243
FT DOMAIN 256 330
FT DISULFID 256 293
FT DISULFID 273 307
FT DISULFID 287 325
FT DISULFID 292 329
FT CONFLICT 28 28
FT CONFLICT 58 58
FT CONFLICT 88 88
FT CONFLICT 103 104
FT CONFLICT 166 169
FT CONFLICT 184 184
FT CONFLICT 200 200
FT CONFLICT 209 209
FT CONFLICT 269 269
FT CONFLICT 284 284
SQ SEQUENCE 349 AA; 37924 MW; 5FFC8EE83EFB4F99 CRC64;

Query Match 39.4%; Score 568; DB 1; Length 349;
Best Local Similarity 46.0%; Pred. No. 3.3e-36;
Matches 110; Conservative 30; Mismatches 81; Indels 18; Gaps 5;

OY 13 SFCLLSM-----VCAQLCRTPTCTPMT-PQCPQGVPLVLDGCGCCVCARRLGESCDHL 67
DB 12 AFVLLALCSRPASSQDCSAPCQCPAGAPRCPAGVSLVLDGCGCCRCVCAKQLSLCTER 71
OY 68 HVCDPGQGLVCPGAGPGGHGAVCLLDEDDGSCSEVNGRRYLDGETFKPNCRVLCRDDG 127
DB 72 DPCDPHKGFLCDFGSPANKIGVCTA-KDGAFCVFGGTVYQSGESFQSSCKYQCTCLDGS 130
OY 128 FTCLPLCSEVDVRLPSWDCPRPKRIQVPGKCCPEWVCDOGVTPAIQRTAQGHQSLA--- 184
DB 131 VGCVPPLCSVDVRLPSDPCPFPRRYKLPKCCCEWVCDEP-----KEHTVGPALAAIRPE 185
OY 185 ----VTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNRFCQLEIQRLCLPRPCLAA 239
DB 186 DTFGPDPMTIRANCLVOTTEWSACSKTCGMGISTRTVNDNAFCRLKQSRKLCMVPCPEA 244

RESULT 5
CTGF_HUMAN STANDARD; PRT; 349 AA.
ID CTGF_HUMAN STANDARD; PRT; 349 AA.
AC P29279; Q96QX2;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor (Hypertrophic chondrocyte-
DE specific protein 24).
GN CTGF OR HCS24.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
```


Accession	Source	Sequence
GN	CTGF.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=20145935; PubMed=10679821;	
RA	Xu J., Smock S.L., Safadi F.F., Rosenzweig A.B., Odgren P.R.,	
RA	Marks S.C. Jr., Owen T.A., Popoff S.N.;	
RT	"Cloning the full-length cDNA for rat connective tissue growth factor:	
RT	implications for skeletal development.";	
RL	J. Cell. Biochem. 77:103-115(2000).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RA	Tezuka K., Tamatani T.;	
RT	"Rattus norvegicus connective tissue growth factor.";	
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.	
CC	-1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY	
CC	VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND	
CC	DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL	
CC	ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA	
CC	SYNTHESIS (BY SIMILARITY).	
CC	-1- SUBCELLULAR LOCATION: Secreted.	
CC	-1- SIMILARITY: Contains 1 IGFBP domain.	
CC	-1- SIMILARITY: Contains 1 WFEC domain.	
CC	-1- SIMILARITY: Contains 1 TSP type-1 domain.	
CC	-1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).	
CC	-----	
DR	EMBL; AF120275; AAD39132.1; -	
DR	EMBL; AB023068; BAA82125.1; -	
DR	InterPro: IPR006208; Cys_knot.	
DR	InterPro: IPR006207; Cys_knot_C.	
DR	InterPro: IPR000867; Insl_gro_fac_pr.	
DR	InterPro: IPR000884; TSP1.	
DR	InterPro: IPR001007; WFEC.	
DR	Pfam; PF00007; Cys_knot; 1.	
DR	Pfam; PF00219; IGFBP; 1.	
DR	Pfam; PF00090; tsp_1; 1.	
DR	Pfam; PF00093; vwc; 1.	
DR	SMART; SM00041; CT; 1.	
DR	SMART; SM00121; IB; 1.	
DR	SMART; SM00209; TSP1; 1.	
DR	SMART; SM00214; vwc; 1.	
DR	PROSITE; PS01185; CTCK_1; 1.	
DR	PROSITE; PS01225; CTCK_2; 1.	
DR	PROSITE; PS00222; IGF_BINDING; 1.	
DR	PROSITE; PS50092; TSP1; 1.	
DR	PROSITE; PS01208; WFEC_1; 1.	
DR	PROSITE; PS50184; WFEC_2; 1.	
KW	Cell adhesion; DNA synthesis; Extracellular matrix; Signal.	
FT	SIGNAL	1 24
FT	CHAIN	25 347
FT	DOMAIN	31 98
FT	DOMAIN	99 165
FT	DOMAIN	196 241
FT	DOMAIN	254 328
FT	DISULFID	254 291
FT	DISULFID	271 305
FT	DISULFID	282 321
FT	DISULFID	285 323
FT	DISULFID	290 327
FT	CONFLICT	35 35
FT	CONFLICT	94 94
FT	SEQUENCE	347 AA; 37756 MW; CFBELAI976B7B16 CRC64;

```

Query Match          38.2%; Score 550.5; DB 1; Length 347;
Best Local Similarity 45.4%; Pred. No. 6.9e-35;
Matches 108; Conservative 28; Mismatches 83; Indels 19; Gaps 5;

QY      15 LCLLSMVC-----AQLCRTPCTC-PWTPPOCPQGVPLVLDDCGCCCKVCARRLGESCDELH 68
      11 : : : | | | | | : : : | | : : : | | | | | : : : | | | | |
Db      11 LALVLLCTRPATGQDCSAQCCQCAAEAPRCAGVSLVLDGCGCCRCVCAKQLGELCTERD 70

QY      69 VCDPSOGLVCQPGAGPGHGAVCLLDEDDGSCENVNGRRYLDGETEFPNCRVLCRCDGCF 128
      11 : : : | | | | | : : : | | : : : | | | | | : : : | | | | |
Db      71 PCDPHKGLCFDFGSPANRKGIVCTA-KDGAPCVFEGSVYRSGESFOSSCKYQCTCLDGAV 129

QY      129 TCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVWCDQGYTPAIQRTAQHQLSAL----- 184
      11 : : : | | | | | : : : | | : : : | | | | | : : : | | | | |
Db      130 GCVPLCSMDVRLPSPDCFPFRRVKLPKCCCEEWVCDPE-----KDRTVVGPALAAVRLLED 184

QY      185 ---VTPASADAPCPNMTAWGCPSTTCGLGIATRVSNONRFCQLEIQRRLLCLPRPCLA 239
      11 : : : | | | | | : : : | | : : : | | | | | : : : | | | | |
Db      185 TFGPDPTMRANCLVQTTTMSACSKTCGMGISTRTVNTDFCRLEKQSRLCMVRPCEA 242

RESULT 7
CTGF_MOUSE
ID      CTGF_MOUSE      STANDARD;      PRT;      348 AA.
AC      P29268; Q922U0;
DT      01-DEC-1992 (Rel. 24, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Connective tissue growth factor precursor (FISP-12 protein)
DE      (Hypertrophic chondrocyte-specific protein 24).
DE      CTGF OR FISP12 OR FISP-12 OR HCS24.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91363290; PubMed=1888698;
RA      Rysceck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;
RT      "Structure, mapping, and expression of flsp-12, a growth factor-
RT      inducible gene encoding a secreted cysteine-rich protein.";
RL      Cell Growth Differ. 2:225-233(1991).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91229699; PubMed=2029337;
RA      Brunner A., Chinn J., Neubauer M.G., Purchio A.F.;
RT      "Identification of a gene family regulated by transforming growth
RT      factor-beta.";
RL      DNA Cell Biol. 10:293-300(1991).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```



```

RN [4]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE=97327410; Pubmed=9184077;
RA Kireeva M.L., Latinkic B.V., Kolesnikova T.V., Chen C.C., Yang G.P.,
RA Abler A.S., Lau L.F.;
RT "Cyr61 and Fisp12 are both ECM-associated signaling molecules:
RT activities, metabolism, and localization during development.";
RL Exp. Cell Res. 233:63-77(1997).
RN [5]
RP FUNCTION.
RX MEDLINE=99182484; Pubmed=10082563;
RA Babic A.M., Chen C.C., Lau L.F.;
RT "Fisp12/mouse connective tissue growth factor mediates endothelial
RT cell adhesion and migration through integrin alphavbeta3, promotes
RT endothelial cell survival, and induces angiogenesis in vivo.";
RL Mol. Cell. Biol. 19:2958-2966(1999).
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM.
CC -1- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN
CC (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
CC -1- INDUCTION: By growth factors.
CC -1- SIMILARITY: Contains 1 IGF1P domain.
CC -1- SIMILARITY: Contains 1 VMF domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M70641; AAA37627.1; -.
DR EMBL; M70642; AAA37628.1; -.
DR EMBL; M80263; AAA73135.1; -.
DR EMBL; BC006783; AAH06783.1; -.
DR PIR; A40578; A40578.
DR MGD; MGI:95537; Ctgf.
DR GO; GO:0005578; C:extracellular matrix; IDA.
DR GO; GO:0008201; F:heparin binding activity; IDA.
DR GO; GO:0005178; F:integrin binding activity; IDA.
DR GO; GO:0001525; P:angiogenesis; IDA.
DR GO; GO:0016477; P:cell migration; IDA.
DR GO; GO:0007160; P:cell-matrix adhesion; IDA.
DR GO; GO:0008543; P:FGF receptor signaling pathway; IDA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IDA.
DR GO; GO:0001503; P:ossification; IMP.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Ins1_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VMF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGF1P; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vmc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VMFC_1; 1.

```

[illegible]

```

CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 VWFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59284; CAA41975.1; -
DR PIR: S20078; S20078.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF00093; vwc; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS0185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; VWFC_1; 1.
DR PROSITE: PS50184; VWFC_2; 1.
DR PROSITE: PS50184; VWFC_2; 1.
KW Proto-oncogene; Growth factor; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 351 NOV PROTEIN.
FT DOMAIN 31 103 IGFBP.
FT DOMAIN 104 170 VWFC.
FT DOMAIN 201 246 TSP TYPE-1.
FT DOMAIN 258 332 CTCK.
FT DISULFID 258 295 BY SIMILARITY.
FT DISULFID 275 309 BY SIMILARITY.
FT DISULFID 286 325 BY SIMILARITY.
FT DISULFID 289 327 BY SIMILARITY.
FT DISULFID 294 331 BY SIMILARITY.
FT CARBOHYD 274 274 N-LINKED (GLCNAC. ) (POTENTIAL).
SQ SEQUENCE 351 AA; 38268 MW; 1ECB3FA3058C6797 CRC64;

Query Match 37.6%; Score 541.5; DB 1; Length 351;
Best Local Similarity 43.6%; Pred. No. 3.3e-34;
Matches 109; Conservative 34; Mismatches 82; Indels 25; Gaps 7;

OY 2 RGSPLIHLATSFCLLSMVC-----AQLCTPC--TCPWTPPQCPQGVPLVLDGCGCK 54
DB 7 QGLPVLL-----LLLLRPCEVSGREACPRPCGRCAPAEPPRCAGVPAVLDCGCGCL 61
OY 55 VCARRLGESCDHLHVCDPSQGLVCQPGAGPGHGAVCLLDEDDGCEVNGRRYLDGETFK 114
DB 62 VCARQGESCSPLLPDESGLYCDRPGEDGGAGICMVLGD-NCVFDGMIYRNGETFQ 120
OY 115 PNCRVLCRDGCGFTCLPLCSEVRLPSPWCPRPKRIQVPGKCCPEWVCDQ-----GV 167
DB 121 PSCKYQCTCRDGOIGCLPRCNLGLLPBDCPFPRIKIEVPGECEKWCDDPRDEVLLGCF 180
OY 168 TPAIQSTAGHQLSALVTPASADAPCPWSTAWGPCSTTGGLGATATRVSNQNRFCQLEI 227
DB 181 AMAAYRQEA---TLGIDVSDSSAN--CIEOTTEWSACSKSGGMSFSTRVYTRNQCEMVK 235
OY 228 QRRICLPPRC 237
DB 236 QTRLCMRPC 245

```

```

RESULT 9
CE10_CHICK
ID CE10_CHICK STANDARD; PRT; 375 AA.
AC P19336;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CER-10 protein precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89145206; PubMed=2537491;
RA Simmons D.L., Levy D.B., Yannoni Y., Erikson R.L.;
RT "Identification of a phorbol ester-repressible v-src-inducible gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:1178-1182(1989).
CC -1- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- INDUCTION: BY V-SRC.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 VWFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J04496; AAA48661.1; -
DR PIR: A41428; A41428.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF00093; vwc; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS0185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; VWFC_1; 1.
DR PROSITE: PS50184; VWFC_2; 1.
KW Growth factor binding; Signal.
FT SIGNAL 1 22
FT CHAIN 23 375 CER-10 PROTEIN.
FT DOMAIN 98 164 VWFC.
FT DOMAIN 223 268 TSP TYPE-1.
FT DOMAIN 281 355 CTCK.
FT DISULFID 281 318 BY SIMILARITY.
FT DISULFID 298 332 BY SIMILARITY.
FT DISULFID 309 348 BY SIMILARITY.
FT DISULFID 312 350 BY SIMILARITY.
FT DISULFID 317 354 BY SIMILARITY.
SQ SEQUENCE 375 AA; 40651 MW; 95F28553BE35D5AE CRC64;

Query Match 37.6%; Score 541; DB 1; Length 375;
Best Local Similarity 38.4%; Pred. No. 3.9e-34;
Matches 99; Conservative 35; Mismatches 94; Indels 30; Gaps 3;

```

QY	10	LATSFLCLLSMCAQLCRTPTCTPWTTPPOCCPQGVPLVLIDGCCGCKVCARRIGESC	DHLHV	69
Db	10	LAALLLCLARLALGSPCPAVCCCPAAAPQCAPGVGLVPDGGCGCKVCARQLNEDCS	RTQP	69
QY	70	CDPSQGLVCQPGAGPGGHGAVCLTDEDDGSCVEVNGRRYLDETFFKPNCRVLC	RCDG	129
Db	70	CDHTKGLFCNFGASPAATNGICRAQSEGRPCCEYNSKIYONGESFQPNCKHQ	CTCIDGAVG	129
QY	130	CLPLCSEDVRLPMSWDCPRPKRIQVPGKCCPEWVCDQGV-----	TPAIQRS	174
Db	130	CIPLCPOELSLPNTLGCPSRLVAYFGQCCCEWVCDESKDALEELLEGFSKE	FGLDASEGE	189
QY	175	TAQGHQLSALV-----TPASADAP-----	CPNWSTAWGPCSTGCLGIATRVSQ	219
Db	190	LTRNNELIAIVKGLKMLPVFGSEFQSRAFENPKCIYQTTWSQCSKTG	TGISTRTVND	249
QY	220	NRFCQLEIQRLCLPRPC	237	
Db	250	NPDKCLIKETRICEVRPC	267	

RESULT	10		
CTGF_PIG			
ID	CTGF_PIG	STANDARD:	PRT; 349 AA.
AC	019113;		
DT	15-JUL-1998	(Rel. 36, Created)	
DT	15-JUL-1998	(Rel. 36, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	Connective tissue growth factor precursor.		
GN	CTGF.		

OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Uterus;
RX MEDLINE=97390475; PubMed=9242708;
RA Brigstock D.R., Steffen C.L., Kim G.Y., Vegunta R.K., Diehl J.R.,
RA Harding P.A.;
RT "Purification and characterization of novel heparin-binding growth
RT factors in uterine secretory fluids. Identification as heparin-
RT regulated Mr 10,000 forms of connective tissue growth factor.";
RL J. Biol. Chem. 272:20275-20282(1997).
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGF1P domain.
CC -1- SIMILARITY: Contains 1 WFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
EMBL; U83916; AAC48756.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.

DR	pfam;	PF00219;	IGFBP;	1.
DR	pfam;	PF00090;	tsp_1;	1.
DR	pfam;	PF00093;	vwf;	1.
DR	SMART;	SM00041;	CT;	1.
DR	SMART;	SM00121;	IB;	1.
DR	SMART;	SM00209;	TSP1;	1.
DR	SMART;	SM00214;	vwf;	1.
DR	PROSITE;	PS01185;	CTCK_1;	1.
DR	PROSITE;	PS01225;	CTCK_2;	1.
DR	PROSITE;	PS00222;	IGF_BINDING;	1.
DR	PROSITE;	PS50092;	TSP1;	1.
DR	PROSITE;	PS01208;	vwf_1;	1.
DR	PROSITE;	PS50184;	vwf_2;	1.
KW	Cell. adhesion;	DNA synthesis;	Extracellular matrix;	Signal.
FT	SIGNAL	1	26	POTENTIAL.
FT	CHAIN	27	349	CONNECTIVE TISSUE GROWTH FACTOR.
FT	DOMAIN	33	100	IGFBP.
FT	DOMAIN	101	167	vwf.
FT	DOMAIN	198	243	TSP TYPE-1.
FT	DOMAIN	256	330	CTCK.
FT	DISULFID	256	293	BY SIMILARITY.
FT	DISULFID	273	307	BY SIMILARITY.
FT	DISULFID	284	323	BY SIMILARITY.
FT	DISULFID	287	325	BY SIMILARITY.
FT	DISULFID	292	329	BY SIMILARITY.
SQ	SEQUENCE	349 AA;	38007 MW;	BB510E2B2B52D4A0 CRC64;

Query Match	37.0%;	Score 533;	DB 1;	Length 349;
Best Local Similarity	44.4%;	Pred. No. 1.5e-33;		
Matches 106;	Conservative 30;	Mismatches 85;	Indels 18;	Gaps 5;

[illegible]

```

RESULT 11
NOV_COTJ A
ID NOV_COTJ A STANDARD; PRT; 353 AA.
AC P42642;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
GN NOV.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria.; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBL_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RA Weiskirchen R., Bister K.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC GROWTH REGULATION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 VWFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:32:15 ; Search time 96 Seconds

(without alignments)
672.012 Million cell updates/sec

Title: US-10-010-408-2

Perfect score: 1440
Sequence: 1 MRGSPLHLATSFCLLSM.....LCIPRCLARSHSWNSAF 250

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 segs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_ricent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1302.5	90.5	251	11	Q8CIC8	Q8cic8 mus musculu
2	575.5	40.0	347	13	Q9PT80	Q9pt80 notophthalm
3	551	38.3	344	13	Q98TQ8	Q98tq8 gallus gall
4	550.5	38.2	343	13	Q42607	Q42607 xenopus lae
5	529.5	36.8	354	11	Q8CA67	Q8ca67 mus musculu
6	517.5	35.9	375	13	Q98TX5	Q98tx5 xenopus lae
7	505	35.1	349	6	O97765	O97765 sus scrofa
8	503.5	35.0	367	11	Q99PP0	Q99pp0 rattus norv
9	503.5	35.0	367	11	O54775	O54775 mus musculu
10	503	34.9	379	11	Q9WTM9	Q9wtm9 rattus norv
11	498.5	34.6	367	4	O95388	O95388 homo sapien
12	455	31.6	176	13	Q9PSS6	Q9ps6 gallus gall
13	372	25.8	331	4	O95958	O95958 homo sapien
14	371.5	25.8	354	4	O95389	O95389 homo sapien
15	355.5	24.7	334	4	Q9UID7	Q9uid7 homo sapien
16	303	21.0	128	11	Q9R2C0	Q9r2c0 rattus norv

17	294.5	20.5	100	4	Q9UDL6	Q9udl6 homo sapien
18	291	20.2	280	4	Q9HCS3	Q9hcs3 homo sapien
19	257	17.8	230	4	Q8WYK7	Q8wyk7 homo sapien
20	210	14.6	374	5	Q9VVK0	Q9vvk0 drosophila
21	210	14.6	470	5	Q9VVK3	Q9vvk3 drosophila
22	200	13.9	77	4	Q9UDE4	Q9ude4 homo sapien
23	161.5	11.2	1664	5	Q9TVO2	Q9tvq2 caenorhabdi
24	161	11.2	626	4	Q8ND91	Q8nd91 homo sapien
25	161	11.2	969	4	Q96KG6	Q96kg6 homo sapien
26	158.5	11.0	1637	6	Q9XSV8	Q9xsv8 bos taurus
27	158.5	11.0	5146	6	Q8SPM4	Q8spm4 bos taurus
28	156.5	10.9	4123	4	O75851	O75851 homo sapien
29	154.5	10.7	747	11	Q8VHF4	Q8vhf4 mus musculu
30	154.5	10.7	1034	11	Q8VHL7	Q8vhl7 mus musculu
31	153.5	10.7	1004	11	Q8CGA7	Q8cga7 mus musculu
32	153.5	10.7	1034	11	Q8VIK5	Q8vik5 mus musculu
33	150.5	10.5	1048	13	Q8AWW5	Q8aww5 gallus gall
34	149.5	10.4	2327	13	Q9IBG7	Q9ibg7 xenopus lae
35	149	10.3	1028	11	Q9JLL0	Q9jll0 mus musculu
36	148.5	10.3	947	11	Q8BKK7	Q8bkk7 mus musculu
37	144	10.0	685	6	Q9TTS5	Q9ts5 bos taurus
38	144	10.0	1246	4	O75095	O75095 homo sapien
39	139.5	9.7	406	11	Q925I3	Q925i3 mus musculu
40	139.5	9.7	426	11	Q8VEA6	Q8vea6 mus musculu
41	139.5	9.7	482	11	Q9CVG8	Q9cvg8 mus musculu
42	139.5	9.7	483	11	Q8KQ1	Q8kq1 mus musculu
43	139.5	9.7	938	11	Q8CIA2	Q8cia2 mus musculu
44	139	9.7	425	6	O02661	O02661 bos taurus
45	139	9.7	507	5	Q8MS81	Q8ms81 drosophila

ALIGNMENTS

RESULT 1	
Q8CIC8	PRELIMINARY; PRT; 251 AA.
ID	Q8CIC8
AC	Q8CIC8;
DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-MAR-2003 (TREMBLrel. 23, last sequence update)
DT	01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE	Similar to WNT1 inducible signaling pathway protein 2.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA	Strausberg.R.;
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC032877; AAH32877.1; -
SO	SEQUENCE 251 AA; 27106 MW; 57520309A9069524 CRC64;

Query Match	90.5%; Score 1302.5; DB 11; Length 251;
Best Local Similarity	89.6%; Pred. No. 2.2e-117;
Matches	225; Conservative 9; Mismatches 16; Indels 1; Gaps 1;
QY	1 MRGSPLHLATSFCLLSMVCACQLCRTPCGPMWPQCPQGVPLVLDGCGCKVCARRL 60
DB	1 MRGNPLHLAIFSLCLLSMVAQAQLCPACACPMTPPQCPGVPLVLDGCGCCARCARL 60
QY	61 GESCDHLHVCPSQGLVCQPGAGPGGHGAVCLDEDDGSCSEVNGRRYLDGETFKPNCRYL 120
DB	61 GESCDHLHVCNPSQGLVCQPGAGPSGRGVCLFEEDDGSCSEVNGRRYLDGETFKPNCRYL 120
QY	121 CRCDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCQGV-TPAIQRTAQGH 179
DB	121 CRCDDGGFTCLPLCSEDVRLPSWDCPRPKRIQVPGKCCPEWVCQGV-TPAIQRTAQGH 180
QY	180 QLSALVTPASADAPCPNMSTAWGPCSTTCGLGATRVSNQNRFCQLEIQRRLCLPRCLA 239
DB	181 QLSALVTPASADAPCPNMSTAWGPCSTTCGLGATRVSNQNRFCQLEIQRRLCLSRPCLA 240


```
OY 240 ARSHSSWNSAF 250
DB 241 SRSHGWSNSAF 251

RESULT 2
O9PT80 PRELIMINARY; PRT; 347 AA.
AC 09PT80;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Connective tissue growth factor.
GN CTGF.
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC Notophthalmus.
OC NCBI_TaxID=8316;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Forelimb blastema;
RA Gates P.B.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Forelimb blastema;
RX MEDLINE=99033008; PubMed=9813273;
RA Cash D.E., Gates P.B., Imokawa Y., Brookes J.P.;
RT "Identification of new connective tissue growth factor as a target of
RT retinoid regulation in limb blastenal cells.";
RL Gene 222:119-124(1998).
DR EMBL; AJ271167; CAB65965.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac-pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWF_C; 1.
SQ SEQUENCE 347 AA; 38098 MW; 3B7E2399F27672C1 GRC64;

Query Match 40.0%; Score 575.5; DB 13; Length 347;
Best Local Similarity 47.5%; Pred. No. 2.6e-47;
Matches 113; Conservative 26; Mismatches 84; Indels 15; Gaps 4;

OY 9 LATSFCLLSMVCQAQLCRTPCTCPWTPPQCPOGVPLVLDGCGCCCKVCARRLGESCDHLH 68
DB 13 LLAVALLSMVS--CAQDCSGECRCPCPKRPECPAGTSLVMDGCGCCCKYCAKOLGELCTEKD 70
OY 69 VCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCCEVNGRRYLDGETFRPNCRVLCRDDGGF 128
DB 71 VCDPHRGLFCDFGSRVNNKIGVCTA-KDGAFCVFGGMVYRSGESFQSSCKYQCTCLDGGV 129
OY 129 TCLPLCSEVRLPSWDCPRPKRIQVPGKCCPEWVCDQ-----GVTPAIQSTAGHQH 181
DB 130 GCVPLCGVDRLPSPDCPFPRRVKQLGKCCCEWVCDQPKQOTRVGPALAVYRQEEYGPD 189
OY 182 SALVTPASADAPCPNMWSTAWGPCSTGCGIATRVSNQNRFCOLEIQRLCLPRCLA 239
DB 190 SSLMR-----ANCLVOTTEWSACSKTCGMGISTRTVNDNEMCRLEKQSRRLCMVRPCEA 242
```

```
RESULT 3
O98TQ8 PRELIMINARY; PRT; 344 AA.
AC 098TQ8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Connective tissue growth factor precursor (Connective tissue growth
DE factor/hypertrophic chondrocyte-specific protein 24).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Gygi D.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Mukudai Y., Kubota S., Takigawa M.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ298335; CAC33438.1; -
DR EMBL; AF463517; AAL68834.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac-pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWF_C; 1.
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 344 CONNECTIVE TISSUE GROWTH FACTOR.
SQ SEQUENCE 344 AA; 37499 MW; 69E639AF69BF1D00 CRC64;

Query Match 38.3%; Score 551; DB 13; Length 344;
Best Local Similarity 45.5%; Pred. No. 5.8e-45;
Matches 111; Conservative 25; Mismatches 94; Indels 14; Gaps 4;

OY 4 SPLIHLATSFCLLSMVCQAQLCRTPCTCPWTP-PQCPQGVPLVLDGCGCCCKVCARRLGE 62
DB 2 SPASLAVALLALLGPEVRGQECSCGCGSGPGSPCAGVSLVLDGCGCCRVCAKOLGE 61
OY 63 SCDHLHVCDSOGLVCPGAGPGGHGAVCLLDEDDGSCCEVNGRRYLDGETFRPNCRVLCR 122
DB 62 LCTERDPCDHHKGLFCDFGSPANRRIGVCTA-RDGAFCVFSGMVYRSGESFQSSCKYQCT 120
OY 123 CDDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGVTPAIQSTAGHQHLS 182
DB 121 CLDGAVGCVPLCSMDVRLPSDCPFPRRVKQLPGKCCCEWVCDQ-----KEQTAVGPALA 175
OY 183 AL-----VTPASADAPCPNMWSTAWGPCSTGCGIATRVSNQNRFCOLEIQRLCLPR 235
DB 176 AYRLEDITYGPDPTMRANCLVQTTESACSKTCGMGISTRTVNDNAFCRLEKQSRRLCMVR 235
OY 236 PCLA 239
DB 236 PCEA 239
```

```

RESULT 4
042607
ID      O42607          PRELIMINARY;           PRT;         343 AA.
AC      O42607;
DT      01-JAN-1998 (TREMBLrel. 05, Created)
DT      01-JAN-1998 (TREMBLrel. 05, last sequence update)
DE      01-MAR-2003 (TREMBLrel. 23, last annotation update)
DEF     Connective tissue growth factor XCTGF.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      SEQUENCE FROM N.A.
RL      Ying Z., King M.L.;
RA      Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; U43524; AAB67639.1; -.
DR      EMBL; U43523; AAB67638.1; -.
DR      InterPro; IPR006208; Cys_knot.
DR      InterPro; IPR006207; Cys_knot_C.
DR      InterPro; IPR000867; Insl_gro_fac_pr.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR001007; VWF_C.
DR      Pfam; PF00007; Cys_knot; 1.
DR      Pfam; PF00219; IGFBP; 1.
DR      Pfam; PF00090; tsp_1; 1.
DR      Pfam; PF00093; vwc; 1.
DR      SMART; SM00041; CT; 1.
DR      SMART; SM00121; IB; 1.
DR      SMART; SM00209; TSP1; 1.
DR      SMART; SM00214; VWC; 1.
DR      PROSITE; PS01185; CTCK_1; 1.
DR      PROSITE; PS01225; CTCK_2; 1.
DR      PROSITE; PS50092; TSP1; 1.
DR      PROSITE; PS01208; VWFC; 1.
SQ      SEQUENCE   343 AA;  37966 MW;  93FE221C5DB565A81 CRC64;

Query Match               38.2%; Score 550.5; DB 13; Length 343;
Best Local Similarity    45.7%; Pred. No. 6.5e-45;
Matches 107; Conservative 26; Mismatches 84; Indels 17; Gaps 4;

OY      17 LLSMVC----AQLCRPTCTCPWTTPQCQGVPLVLDGCGCCVKCARRLGESCDHLHYCDP 72
       |::|  |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      11 LFALFCWVSDAQECNGEQCCPNKVPVCDPGVRMVQDGGCGCKVKCSKLGLCTERDVCDP 70
OY      73 SGLVCQPGAGPGGHGAVCLLDEDGSCSEVNRRYLDETETFKPNCRVLRCDDGGFTCLP 132
       :|||  |::|  |::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      71 HKGLCFDFGSRVNRKIGCYTARE-GAPCVFGTVYRSGESFQSSCKKYQCTCIDGGVGCV 129
OY      133 LCSDEVRLPSWDCPPRKRIQVPGKCCEPEWVCDQGVTPAIQRSTAQHQLSAL-----V 185
       |||  |::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      130 LCSMDIRLPSEPCEPRPVKVLPGKCCEEWVCDQP-----QERTLVGPALPAFRMEETYGP 184
OY      186 TPASADAPCPNMSTAWGPCSTTCGLGIATRVSNNRRFCQLEIQRLCLPRPCLA 239
       |::|  |::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      185 DPSLIRANCLVQTTEMSACSKTCGMGISTRYTINDNEHCRLEKQSRCLCMVRPCEA 238

RESULT 5
08CA67
ID      08CA67          PRELIMINARY;           PRT;         354 AA.
AC      08CA67;
DT      01-MAR-2003 (TREMBLrel. 23, Created)
DT      01-MAR-2003 (TREMBLrel. 23, last sequence update)
DE      01-MAR-2003 (TREMBLrel. 23, last annotation update)
DEF     Nephroblastoma overexpressed.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=spinal cord;
RA MEDLINE=22354683; PubMed=12466851;
RX The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL: AK039481; BAC30363.1; -.
SQ SEQUENCE 354 AA; 38828 MW; 6F3B34A3A84187EE CRC64;

Query Match      36.8%; Score 529.5; DB 11; Length 354;
Best Local Similarity 42.5%; Pred.No.7e-43;
Matches 105; Conservative 26; Mismatches 87; Indels 29; Gaps 6

QY 10 LATSFLC--LLSMVCAQL---CRTPTCTCPWTPPQCPCGCVPLVDGCCCKVCARRLGESC 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10 LCLGFLFHLLSQVSASLRPCSRCPKPCPSISPTCAPGVRSVLDDGCCCPVCARORGESC 69
   | | | | | | | | | | : | : | : | : | : | : | : | : | : | : | : |

QY 65 DHLHVCDPSSQGLVCCPGAGPGGHGAVCLLDEDDSGCEVNGRRYLDETEFKPNCRVLCRCD 124
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 70 SEMRPCDQSSGLYCDSADENNQTGICMVPEGD-NCVFDGVITYRNGEKEFPNCQYFCTCR 128
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 125 DGGFTCLPLCSSEDRVLRLPSPWDPCRPRKRIQVP GKCCPEWVDCDGVTVAIQRTAQGHQ--LS 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 129 DGQIGCLPRCQLDVLLPGPDPCAPARKVAVPGECCEKWTCG-----SDEQGTQGTILG 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 183 ALVTPA-----SADAPCPNMSTAWGPCSTTCGLGIATRVSNQRFCQLEIQR 230
   -||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
DB 180 GLALPAYRPEATVGVEVSDSIINCIEQTETWSACSCKSCGMGVSTRVTNRNRCQCEMVKQTR 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 231 LCLPRPC 237
   ||| |||
DB 240 LCIVRPC 246

RESULT 6
Q98TX5 PRELIMINARY; PRT; 375 AA.
ID Q98TX5 AC Q98TX5:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Secreted cysteine-rich protein cyr61.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Latinkic B.V., Bennett B., Smith J.C.;
RT "Characterization of Xenopus cyr61.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF320592; AAK00947.1; -.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSPI.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF00093; vwc; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSPI; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSPI; 1.
DR PROSITE: PS01208; VWFC; 1.
```

SO SEQUENCE 375 AA; 41460 MW; 78075CA7B380304E CRC64;
Query Match 35.98; Score 517.5; DB 13; Length 375;
Best Local Similarity 37.2%; Pred. No. 1.1e-41;
Matches 99; Conservative 39; Mismatches 87; Indels 41; Gaps 5;
QY 13 SFLLLSMVCACL-----CRPTCTCPWTPPQCPQGVPLVLDGCGCCCKVCARRLG 61
DB 2 SFLLNPLVALIALISGFIDLAVSSCPAVCCPEVPPKCAPGVGLVLDGCGCCCKICAKOLN 61
QY 62 ESCDHLHVCDPGSLVCPGAGPGGHGAVCLDEDDGSCVENGRRLDGETFKPNCRVLC 121
DB 62 EDCSKTHPCDHTGIECNFEGASSRAIKGICRAKSEGRPCCEYNISKIYNGESFQPNCKHOC 121
QY 122 RCDDGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGVTPAIO----- 172
DB 122 TCIDGAVGCLPLCPQLSLPLNLCGPNRLVYVGCCCEWVCDCAKDPVDEMDFENKEF 181
QY 173 -RSTAQ-----HQLSALV-----TPAS--ADAPCPNMSTAMGPCSTTCGLG 211
DB 182 GMDTNEGELTRKNEFVAVIKGLKMLPVFGSDPSQSHVENSCKIVQITSMSCSKTCGTG 241
QY 212 IATRVSNQNRFCOLEIQRRLCLPRPC 237
DB 242 ISTRVTNDNSNCRVLVRETRICEVRPC 267
RESULT 7
097765 PRELIMINARY; PRT; 349 AA.
ID 097765;
AC 097765;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE Connective tissue growth factor.
GN CTGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA Harding P.A., Brigstock D.R.;
RT "Cloning and sequencing of a porcine connective tissue growth factor
RT (CTGF) cDNA."
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U70060; AAD00174.1; -.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF00093; vwc; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; vwc; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; VWF; 1.
SQ SEQUENCE 349 AA; 37946 MW; 35AB4275AC1D4B3A CRC64;
Query Match 35.1%; Score 505; DB 6; Length 349;
Best Local Similarity 42.3%; Pred. No. 1.6e-40;
Matches 101; Conservative 32; Mismatches 88; Indels 18; Gaps 5;
QY 13 SFLLLSM-----VCAQLCRTPTCTCPWTPPQ-CPQGVPLVLDGCGCCCKVCARRLGESCDHL 67

DB 12 AFVLLIALCSRPASGQDCSGQCACAGKRRACPAGVSSLLLEGCGCCRCIAKHLGLDCTER 71
QY 68 HVCDPGSLVCPGAGPGGHGAVCLDEDDGSCVENGRRLDGETFKPNCRVLCRDDG 127
DB 72 APCDPHKGLFCDEGSPANRKIGVCTA-KDCAPCFVGGTVYRSGESFGSSCKYQCTCLDGA 130
QY 128 FTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCDQGVTPAIOGSTAOGHLSAL--- 184
DB 131 VGCVPPLCSMDVRLPSPDCPFPRRVKLPGKCCCEWVCDPE-----KHRTVGPALAAVRL 185
QY 185 ----VTPASADAPCPNMSTAMGPCSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRPCLA 239
DB 186 DTFGPDPTMMRANCLVQTEWSACSKTCGMGISTRTVNDNASCRLEKOSRLCMVRPCEA 244

RESULT 8
099PP0 PRELIMINARY; PRT; 367 AA.
ID 099PP0;
AC 099PP0;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE ELM1.
GN ELM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20487548; Pubmed=11031104;
RA Sleeman M.A., Murlison J.G., Strachan L., Kumble K.D., Glenn M.P.,
RA McGrath A., Bickerstaff P., Grierson A., Havukkala I., Tan P.,
RA Watson J.D.;
RT "Gene expression in rat dermal papilla cells: analysis of 2529 ESTs."
RL Genomics 69:214-224(2000).
DR EMBL: AF228049; AAK00729.1; -.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF00093; vwc; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; vwc; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; VWF; 1.
SQ SEQUENCE 367 AA; 40613 MW; 8A4A34C69D3243D2 CRC64;
Query Match 35.0%; Score 503.5; DB 11; Length 367;
Best Local Similarity 43.7%; Pred. No. 2.3e-40;
Matches 94; Conservative 29; Mismatches 89; Indels 3; Gaps 3;
QY 24 QLCRTPTCTCPWTPPQCPQGVPLVLDGCGCCCKVCARRLGESCDHLHVCDPGSLVCPGAG 82
DB 47 EFCKWPCECPQAPRPRCLGVSLITDGECECKICAQQLGDNCTEAAYVCDPHRGLYCDYSGD 106
QY 83 GPGGHGAVCLDEDDGSCVENGRRLDGETFKPNCRVLCRDDGGFTCLPLCSEDEVRLPS 142
DB 107 RPRYAIGVCAQVYGVG-CVLDGVRYTNGESFQPNCRYNCTCIDGIVGCTPLCLSP-RPPR 164
QY 143 WDCPRPKRIQVPGKCCPEWVCDQGVTPAIOGSTAOGHLSALVTPASADAPCPNMSTAMG 202
DB 165 LWCQRPRHVRVPGQCCQWVCDDDARPRRQTAALLDTRAFASGAVGQRYENCIAYTSPWS 224

QY 203 PCSTTCGIGIATRVSNQNRFCQLEIQRRLCLPRPC 237
Db 225 PCSTTCGIGISTRISNVNARCWPEQESRLCNLRPC 259

RESULT 9

054775 PRELIMINARY; PRT; 367 AA.
AC 054775;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE ELM1.
GN ELM1 OR WISPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HEN;
RX MEDLINE=98119879; PubMed=9449709;
RA Hashimoto Y., Shindo-Okada N., Tani M., Nagamachi Y., Takeuchi K.,
Shirolshi T., Toma H., Yokota J.;
RT "Expression of the Elml gene, a novel gene of the CCN (connective
tissue growth factor, Cyr61/Ce10, and neuroblastoma overexpressed
RT gene) family, suppresses in vivo tumor growth and metastasis of K-1735
RT murine melanoma cells."
RL J. Exp. Med. 187:289-296(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=99061933; PubMed=9843955;
RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
Gurney A.L., Botstein D., Levine A.J.;
RT "WISP genes are members of the connective tissue growth factor family
RT that are up-regulated in wt-1-transformed cells and aberrantly
RT expressed in human colon tumors."
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
DR EMBL; AB004873; BAA24949.1; -.
DR EMBL; AF100777; AAC96319.1; -.
DR MGI; 1197008; Wispl.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWF_C; 1.
SQ SEQUENCE 367 AA; 40702 MW; 3B7C0569EFAB5E96 CRC64;

Query Match 35.0%; Score 503.5; DB 11; Length 367;
Best Local Similarity 43.3%; Pred. No. 2.3e-40;
Matches 93; Conservative 31; Mismatches 88; Indels 3; Gaps 3;

QY 24 QLCRTPTCTPWTTPQCPQGVPLVLDDGGCCCKVCARLGECDHLHYCDPSQGLVCQ-PGA 82
Db 47 EFCKWPCPCFPQSPRCPLGLVSLITDGCCECKICAQQLGDNCTEAICDPHRLGYCDYSGD 106
QY 83 GPGHGAVCLLDEDDGSCCEVNGRRYLDGETFKPNCRVLCRCDDGGFTCLPLCSEDEVRLPS 142

Db 107 RPRYAIGCAQVVG-VCLDGVRYINGESFQPNCRYNCTCIDGTVGCTPLCLSP-RPPR 164
QY 143 WDCPRPKRIQVPGKCCPEWVCDDGYTPAIQRSTAQGHLSALVTPASADAPCPNMSTAWG 202
Db 165 LWCRRQPRHVRVPGQCCQWVCDDARPRRQOTALDTRAFASGAVEQRYENCIAYTSPWS 224
QY 203 PCSTTCGIGIATRVSNQNRFCQLEIQRRLCLPRPC 237
Db 225 PCSTTCGIGISTRISNVNARCWPEQESRLCNLRPC 259

RESULT 10

09WTM9 PRELIMINARY; PRT; 379 AA.
AC 09WTM9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CYR61 precursor.
GN CYR61.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Izm; TISSUE=Aorta;
RA Unoki H., Yonekura H., Furukawa K., Yamamoto H.;
RT "Rat Cyr61 mRNA."
Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB015877; BAA78339.1; -.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWF_C; 1.
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 379 CYR61.
SQ SEQUENCE 379 AA; 41728 MW; D2ABAFD77B84762B CRC64;

Query Match 34.9%; Score 503; DB 11; Length 379;
Best Local Similarity 36.2%; Pred. No. 2.7e-40;
Matches 97; Conservative 36; Mismatches 101; Indels 34; Gaps 3;

QY 4 SPLHLATSFCL-LSMVCAQLCRTPTCTPWTTPQCPQGVPLVLDDGGCCCKVCARLGE 62
Db 3 SSTIKTLAVAVTLLHLTRLALSTCPASCHPLAPKCAPGVGLVDDGGCCCKVCAKQLINE 62
QY 63 SCDHHLVCDPSQGLVCQPGAGPGHGAVCLLDEDDGSCCEVNGRRYLDGETFKPNCRVLCR 122
Db 63 DCSKTQPCDHTKGLGCNFGANSTALKGICRAQSEGRPCENRYONGESFQPNCKHQT 122
QY 123 CDDGGFTCLPLCSEDEVRLPSWDCPRPKRIQVPGKCCPEWVCQ----- 165
Db 123 CIDGAVGCIPLCPQELSLPNLGCPRNRLVYVSGQCCCEWVCDEDSIKDSLDDQDDLGLFD 182
QY 166 -----GVTPAIQRSTAQGHLSALVTPASADA-PCPNWSTAWGPCSTTCG 209
Db 183 ASEVELTRNNELIATGKSSSLKRLPLVEGTEPRVLYNPLHAHGOKCIVQTTWSQCSKSCG 242

OY 210 LGIATRVSNONRRCOLEIQRRLCLPRPC 237
 DB 243 TGISTRTVNDNPECRLVKETRICVPRPC 270

RESULT 11

O95388 PRELIMINARY; PRT; 367 AA.
 AC O95388;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Connective tissue growth factor related protein WISP-1.
 GN WISP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung, and fetal kidney;
 RX MEDLINE=99061933; PubMed=9843955;
 RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
 RA Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
 RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
 RA Gurney A.L., Botstein D., Levine A.J.;
 RT "WISP genes are members of the connective tissue growth factor family
 RT that are up-regulated in wt-1-transformed cells (and aberrantly
 RT expressed in human colon tumors".
 RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Blechschmidt K., Kalaydjieva L., Goodman R., Gresham D., Baas F.,
 RA Jonge R.d., Schilhabel M., Schattevoy R., Dette M., Menzel U.,
 RA Rosenthal A.;
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF100779; AAC96321.1; -
 DR EMBL; AF192304; AAF22341.1; -
 DR Genew; HGNC:12769; WISP1.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR000867; Insl_gro_fac-pr.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGF1BP; 1.
 DR Pfam; PF00090; tsp_1; 1.
 DR Pfam; PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00092; TSP1; 1.
 DR PROSITE; PS01208; VWF; 1.
 SQ SEQUENCE 367 AA; 40331 MW; 9F29CA94D69C0502 CRC64;

Query Match

Best Local Similarity 34.6%; Score 498.5; DB 4; Length 367;
 Matches 96; Conservativity 31; Mismatches 81; Indels 11; Gaps 7;

OY 24 QLCRTPCTCPWTPPQCPQGVPLVLDGCGCKVCARLGESECDHLHYCDPSQGLVCO-PCA 82
 DB 47 QFCRWPCPCSPSPRCPLGSLITDGCCECKMCAQQLGDNCTEAICDPHRLGYCDYSGD 106
 OY 83 GPGHGAVCLLDEDDGSCCEVNGRRYLDGETFKPNCRVLCRCDDGFTCLPLCSEDEVRLPS 142
 DB 107 RPRYAIQVCAQVVG-CLVDGVRVNNQSFQPNCKYNCTQIDGAVGCTPLCLR-VRPPR 164
 OY 143 WDCPRPRRIQVPGKCCPEWVC-DQGVTPAIQRSTAQGHQLSALVTPASADA---PCPNWS 198

DB 165 LMCPPHRRVSIPIGHCEQWVCEDDAKRP---RKTAP-RDTGAFDAVGEVAMHRNCIAYT 220
 OY 169 TAWGPCSTTCGLGIATRVSNONRRCOLEIQRRLCLPRPC 237
 DB 221 SPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPC 259

RESULT 12

O9PSS6 PRELIMINARY; PRT; 176 AA.
 AC O9PSS6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Potential IGF binding protein (Fragments).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92096871; PubMed=1756408;
 RA Martinerie C., Perbal B.;
 RT "Expression of a gene encoding a novel potential IGF binding protein
 RT in human tissues".
 RL C. R. Acad. Sci., III, Sci. Vie 313:345-351(1991).
 DR InterPro; IPR000867; Insl_gro_fac-pr.
 DR InterPro; IPR001007; VWF_C.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS01208; VWF; 1.
 FT NON_TER 1
 FT NON_CONS 77 78
 FT NON_CONS 154 155
 FT NON_TER 176 176
 SQ SEQUENCE 176 AA; 18656 MW; 4FA69FFDB79B1C67 CRC64;

Query Match 31.6%; Score 455; DB 13; Length 176;
 Best Local Similarity 42.1%; Pred. No. 5e-36;
 Matches 83; Conservativity 23; Mismatches 59; Indels 32; Gaps 3;

OY 26 CRTPC--TCPWTPPQCPQGVPLVLDGCGCKVCARLGESECDHLHYCDPSQGLVCPAG 83
 DB 10 CPRPCGGRPAEPRAICAPVAVLDGCGCCLVCARQGESCSPLPCDESGGLYCDRGP 69
 OY 84 PGHGAVCLLDEDDGSCCEVNGRRYLDGETFKPNCRVLCRCDDGFTCLPLCSEDEVRLPSW 143
 DB 70 DGGGAGIC---EGDNCVFDMYRNGETFPQSCKYQCTCRDQIGICLLPRCNLGLLPGP 125
 OY 144 DCPRPRIQVPGKCCPEWVCDDGVTPAIQRSTAQGHQLSALVTPASADAPCPNMSTAWCP 203
 DB 126 DCPFPRIQVPGKCCPEWVCDDGVTPAIQRSTAQGHQLSALVTPASADAPCPNMSTAWCP 159
 OY 204 CSTTCGLGIATRVSNON 220
 DB 160 CSKSCGMGFSTRVTNRN 176

RESULT 13

O95958 PRELIMINARY; PRT; 331 AA.

AC O95958;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE DJ142L7.3 (Connective tissue growth factor (NOV, GIG) like protein)
 DE (Fragment).
 GN DJ142L7.3 OR LIBC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

DR SMART; SM00041; CT; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWF; 1.
DR PROSITE; PS01208; VWF; 1.
SQ SEQUENCE 334 AA; 37246 MW; 9188987A7352E948 CRC64;

Query Match 24.7%; Score 355.5; DB 4; Length 334;
Best Local Similarity 30.3%; Pred. No. 3.6e-26;
Matches 74; Conservative 29; Mismatches 66; Indels 75; Gaps 5;

QY 49 GCGCKVCARRLGESCDHL-HVCDP-----SGLVQDGGAGPGHGAVCLLDEDDGSCE 101
DB 2 GCGTHP-----NLCIHLGHTASPTSYKHHKGLCENEGASSTALKICRAQSEGRPCE 54
QY 102 VNGRRYLDGETFRKPNCRVLCRCDGCGFTCLPLCSEDRVDPSPWDCPRPKRIQVPGKCCPEW 161
DB 55 YNSRIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQLSLPPLGCPNPRLVKVTGQCCCEW 114
QY 162 VCDQ-----GVTPIQR 173
DB 115 VCDEDSIKDPMEDQDGLGKELGFDASEVELTRNNELIANGKSSILKRIPVFGMEPRIRY 174
QY 174 STAQGHQLSALVTPASADAPCPNWSSTAWGPCSTTCGLGATRVSNQNRFCOLEIQRLCL 233
DB 175 NPLQGOK-----CIVQTSWSQCSKTCGTGISTRTVINDNPECRLVKETRICE 221
QY 234 PRPC 237
DB 222 VRPC 225

Search completed: July 25, 2003, 12:37:14
Job time : 98 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:59:01 ; Search time 2092.77 Seconds
(without alignments)
10471.464 Million cell updates/sec

Title: US-10-010-408-3
Perfect score: 753
Sequence: 1 ATGAGGGGCGACGCCACTGAT.....CATGGAACAGTCTTCTTAA 753

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:**
1: gb_ba:**
2: gb_htg:**
3: gb_in:**
4: gb_cm:**
5: gb_ov:**
6: gb_pat:**
7: gb_ph:**
8: gb_pl:**
9: gb_pr:**
10: gb_ro:**
11: gb_sts:**
12: gb_sy:**
13: gb_un:**
14: gb_vi:**
15: em_ba:**
16: em_fun:**
17: em_hum:**
18: em_in:**
19: em_mu:**
20: em_cm:**
21: em_or:**
22: em_ov:**
23: em_pat:**
24: em_ph:**
25: em_pl:**
26: em_ro:**
27: em_sts:**
28: em_un:**
29: em_vi:**
30: em_htg_hum:**
31: em_htg_inv:**
32: em_htg_other:**
33: em_htg_mus:**
34: em_htg_pln:**
35: em_htg_rod:**
36: em_htg_mam:**
37: em_htg_vrt:**
38: em_sy:**
39: em_htgo_hum:**
40: em_htgo_mus:**
41: em_htgo_other:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	732.2	97.2	1741	10 AF259981	AF259981 Rattus no
2	659	87.5	1734	6 AR210324	AR210324 Sequence
3	659	87.5	1734	6 AR210325	AR210325 Sequence
4	659	87.5	1734	10 AF100778	AF100778 Mus muscu
5	654.2	86.9	1739	10 AF126063	AF126063 Mus muscu
6	510.4	67.8	1266	6 AX076919	AX076919 Sequence
7	510.4	67.8	1266	6 AX464186	AX464186 Sequence
8	510.4	67.8	1283	9 AF083500	AF083500 Homo sapi
9	510.4	67.8	1293	6 AR210322	AR210322 Sequence
10	510.4	67.8	1293	6 AR210323	AR210323 Sequence
11	510.4	67.8	1309	9 AF074604	AF074604 Homo sapi
12	510.4	67.8	1427	9 AF100780	AF100780 Homo sapi
13	510.4	67.8	1450	9 BC017782	BC017782 Homo sapi
14	501.2	66.6	738	6 AR210337	AR210337 Sequence
15	496.4	65.9	841	6 AR210338	AR210338 Sequence
16	256.6	34.1	137964	2 AC126895	AC126895 Rattus no
17	223.8	29.7	137964	2 AC126895	AC126895 Rattus no
18	223.8	29.7	226303	2 AC095418	AC095418 Rattus no
19	216.2	28.7	61072	10 AL731698	AL731698 Mouse DNA
20	200.4	26.6	216757	2 AL669906	AL669906 Mus muscu
21	169.8	22.5	107260	9 AL139352	AL139352 Human DNA
22	168	22.3	1585	5 NV1271167	AJ271167 Notophtha
23	163	21.6	2075	6 AR018957	AR018957 Sequence
24	163	21.6	2075	6 AR118595	AR118595 Sequence
25	163	21.6	2075	6 AR130891	AR130891 Sequence
26	163	21.6	2075	6 AX206708	AX206708 Sequence
27	163	21.6	2075	6 I11636	I11636 Sequence 1
28	163	21.6	2075	6 I32210	I32210 Sequence 1
29	163	21.6	2075	9 HUMCONGRO	M92934 Homo sapien
30	163	21.6	2312	9 HSCITGF	X78947 H.sapiens m
31	163	21.6	2998	6 AR119211	AR119211 Sequence
32	163	21.6	2998	6 AR151276	AR151276 Sequence
33	161	21.4	2267	6 AX206706	AX206706 Sequence
34	161	21.4	2267	10 MUSFISP12B	M70642 Mouse FISP-
35	161	21.4	2330	10 MUSTGFB	M80263 Mouse mRNA
36	161	21.4	2334	10 BC006783	BC006783 Mus muscu
37	159.2	21.1	1496	4 SSU83916	U83916 Sus scrofa
38	158	21.0	2338	6 E37595	E37595 Monoclonal
39	158	21.0	2338	10 AB023068	AB023068 Rattus no
40	158	21.0	2345	6 AX402019	AX402019 Sequence
41	158	21.0	2345	10 AF120275	AF120275 Rattus no
42	156.4	20.8	2350	6 AR194011	AR194011 Sequence
43	156	20.7	2190	5 AF463517	AF463517 Gallus ga
44	156	20.7	2288	5 GGA298335	AJ298335 Gallus ga
45	154.8	20.6	2350	6 AR201286	AR201286 Sequence

ALIGNMENTS

RESULT 1
AF259981
LOCUS AF259981 1741 bp mRNA ROD 09-MAY-2000
DEFINITION Rattus norvegicus CCN family protein COP-1 (COP-1) mRNA, complete
cds.
ACCESSION AF259981
VERSION AF259981.1 GI:7739780
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1741)
Zhang, R., Averboukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P.J.,

TITLE Coffey,R.J., Pardee,A.B. and Liang,P.
Identification of rCop-1, a new member of the CN protein family,
as a negative regulator for cell transformation
JOURNAL Mol. Cell. Biol. 18 (10), 6131-6141 (1998)
MEDLINE 98414629
PUBMED 9742130
REFERENCE 2 (bases 1 to 1741)
AUTHORS Liang,P.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer
Center, 649 MRB II, Nashville, TN 37232, USA
FEATURES
source location/Qualifiers
1. .1741
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1. .1741
/gene="Cop-1"
262. .1014
/gene="Cop-1"
/note="secreted protein"
/codon_start=1
/product="CCN family protein COP-1"
/protein_id="AAF6901.1"
/db_xref="GI:7739781"
/translation="MRGSPILRLATSFLCLLSMCAQLCRTPTCPWTPPQCPQGV
LVLDGCGCKVCARLRESEHLHVCEPSQGLVCGPAGBPGHGAVALDEDGDEY
NGRRYLDGETFKPNCRVLCDDGFTCLPLCSQEDVTLPSWDCPRPKRIQVPGKCE
MYCDGVTPIQIRSAQGHQSLALVTPASADAPMPNNSIAMWPCSTCGLGIAIRVSN
QNRFCQLEIQRRLCLPRCLIAARSHSSWNSAF"
BASE COUNT 386 a 491 c 480 g 384 t
ORIGIN
Query Match 97.2%; Score 732.2; DB 10; Length 1741;
Best Local Similarity 98.3%; Pred. No. 2e-168;
Matches 740; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 ATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCCTCTGCTCTCTCATG 60
DB 262 ATGAGGGGAGCCCACTGATCCGCTTCTGGCCACTTCCCTCTGCTCTCTCATG 321
QY 61 GTGTGTGCCAGCTGTGCCGAGACCCCTGTACCTGTCTCTTGACACACCCAGTGCCCA 120
DB 322 GTGTGTGCCAGCTGTGCCGAGACCCCTGTACCTGTCTCTTGACACACCCAGTGCCCA 381
QY 121 CAGGGGGTACCCCTGTGTCTGTGATGGCTGTGCTGTAAAGTGTGTGACAGGAGCTG 180
DB 382 CAGGGGGTACCCCTGTGTCTGTGATGGCTGTGCTGTAAAGTGTGTGACAGGAGCTG 441
QY 181 GGGAGTCTTCGACACCTGATGTCTGCAACCCAGCCAGGCGCTGTTGTGAGCCT 240
DB 442 ACGGAGTCTTCGAAACACCTGATGTCTGCAACCCAGCCAGGCGCTGTTGTGAGCCT 501
QY 241 GGGGAGGCGCTGCGGCGCATGGGCTGTGTCTCTTGATGAGATGACGGTAGCTGT 300
DB 502 GGGGAGGCGCTGCGGCGCATGGGCTGTGTCTCTTGATGAGATGACGGTAGCTGT 561
QY 301 GAGGTGAATGGCCGACGTACCTGATGAGAGACCTTTAAACCAATTGACAGGTCCTG 360
DB 562 GAGGTGAATGGCCGACGTACCTGATGAGAGACCTTTAAACCAATTGACAGGTCCTG 621
QY 361 TGCCGCTGTGATGACGGTGGCTTCACTGCTGCGGCTGTGACGTGAGATGTGCGGCTG 420
DB 622 TGCCGCTGTGATGACGGTGGCTTCACTGCTGCGGCTGTGACGTGAGATGTGACGCTG 681
QY 421 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAAGAAAGTGTGCCCCGAG 480
DB 682 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAAGAAAGTGTGCCCCGAG 741
QY 481 TGGGTATGTGACGAGGAGTGAACGCGGATCCAGCGCTCCACGCGCAAGACACCAA 540
DB 742 TGGGTATGTGACGAGGAGTGAACGCGGATCCAGCGCTCCGCGCGCAAGACACCAA 801
QY 541 CTTTCTGCCCTGTCACTCTGCTCTGTGATGCTCTTGTCCAATTGAGACACAGCC 600

DB 802 CTTTCTGCCCTGTCACTCTGCTCTGTGATGCTCTTGGCCAAATTGAGACACAGCC 861
QY 601 TGGGGCCCTGTCAACACCACTGTGGGCTGGGCATAGCACCCGAGTGTCCAACAGAAC 660
DB 862 TGGGGCCCTGTCAACACCACTGTGGGCTGGGCATAGCACCCGAGTGTCCAACAGAAC 921
QY 661 CGATTCTGCCAAGTGAAGATCCACGCGGCTGTGTCTGCGCCAGACCCCTGCTGGCAGCC 720
DB 922 CGATTCTGCCAAGTGAAGATCCACGCGGCTGTGTCTGCGCCAGACCCCTGCTGGCAGCC 981
QY 721 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 753
DB 982 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 1014

RESULT 2
AR210324 1734 bp DNA linear PAT 20-JUN-2002
LOCUS AR210324
DEFINITION Sequence 17 from patent US 6387657.
ACCESSION AR210324
VERSION AR210324.1 GI:21512525
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.
WISP polypeptides and nucleic acids encoding same
Patent: US 6387657-A 17 14-MAY-2002;
TITLE Location/Qualifiers
JOURNAL
FEATURES
source 1. .1734
/organism="unknown"
BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN

Query Match 87.5%; Score 659; DB 6; Length 1734;
Best Local Similarity 93.0%; Pred. No. 1.6e-150;
Matches 702; Conservative 0; Mismatches 50; Indels 3; Gaps 1;
QY 1 ATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCCTCTGCTCTCTCATG 60
DB 257 ATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCCTCTGCTCTCTCATG 316
QY 61 GTGTGTGCCAGCTGTGCCGAGACCCCTGTACCTGTCTCTTGACACACCCAGTGCCCA 120
DB 317 GTGTATPCCAGCTGTGCCAGACCCCTGTGCTCTCTTGACACACCCAGTGCCCA 376
QY 121 CAGGGGGTACCCCTGTGTCTGTGATGGCTGTGCTGTAAAGTGTGTGACAGGAGCTG 180
DB 377 CCGGGGGTACCCCTGTGTCTGTGATGGCTGTGCTGTAAAGTGTGTGACAGGAGCTG 436
QY 181 GGGAGTCTTCGACACCTGATGTCTGCAACCCAGCCAGGCGCTGTTGTGAGCCT 240
DB 437 GGGAGTCTTCGACACCTGATGTCTGCAACCCAGCCAGGCGCTGTTGTGAGCCT 496
QY 241 GGGGAGGCGCTGCGGCGCATGGGCTGTGTCTCTTGATGAGATGACGGTAGCTGT 300
DB 497 GGGGAGGCGCTGCGGCGCATGGGCTGTGTCTCTTGATGAGATGACGGTAGCTGT 556
QY 301 GAGGTGAATGGCCGACGTACCTGATGAGAGACCTTTAAACCAATTGACAGGTCCTG 360
DB 557 GAGGTGAATGGCCGACGTACCTGATGAGAGACCTTTAAACCAATTGACAGGTCCTG 616
QY 361 TGCCGCTGTGATGACGGTGGCTTCACTGCTGCGGCTGTGACGTGAGATGTGCGGCTG 420
DB 617 TGCCGCTGTGATGACGGTGGCTTCACTGCTGCGGCTGTGACGTGAGATGTGCGGCTG 676
QY 421 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAAGAAAGTGTGCCCCGAG 480
DB 677 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAAGAAAGTGTGCCCCGAG 736

QY	481	TGGGTATGTGACCAAGGAGCTGA--CAACGGGGGATCCAGCGCTCCACGGCGCAAGGACAC	537
Db	737	TGGGTGTGTGACCAGGAGCTGATGCAGCGGCAATCCAGCCCTCTCAGCCCAAGGACAC	796
QY	538	CAACTTCTGTGCCCTGTCTCACTCCTGCCTCTGCTGTATGCTCCTGTGCCAAATTGGAGCACA	597
Db	797	CAACTTCTGTGCCCTGTCTCACTCCTGCATCTGCCGATGGCCCTGTCCAAACTGGAGCACA	856
QY	598	GCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCAATAGCCACCCGAGTGTCCAACCAAG	657
Db	857	GCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCAATAGCCACCCGAGTATCCAACCAAG	916
QY	658	AACCGATTCTGCCAACTGGAGATCCAACGCGCCGCTGTGTCTGCCCCAGACCCTGCCTGGCA	717
Db	917	AACCGATTCTGCCAACTGGAGATCCAACGCGTGCCTGTGTCTGTCAGAACCCCTGCCTGGCA	976
QY	718	GCCAGAGGCCACAGCTCATGGAACAGAGTGCCTTCTA	752
Db	977	TCCAGGAGCCACGCGCTCATGGAACAGAGTGCCTTCTA	1011

RESULT 3					
AR210325/c					
LOCUS	AR210325	1734 bp	DNA	linear	PAT 20-JUN-2002
DEFINITION	Sequence 18 from patent US 6387657.				
ACCESSION	AR210325				
VERSION	AR210325.1	GI:21512526			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				

REFERENCE	1 (bases 1 to 1734)
AUTHORS	Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pernica,D., Roy,M.Ann. and Wood,W.I.
TITLE	WISP polypeptides and nucleic acids encoding same
JOURNAL	Patent: US 6387657-A 18 14-MAY-2002;
FEATURES	Location/Qualifiers
source	1..1734

	/organism="unknown"			
BASE COUNT	393 a	495 c	491 g	355 t
ORIGIN				

Query Match	87.58;	Score 659;	DB 6;	Length 1734;
Best Local Similarity	93.08;	Pred. No. 1.6e-150;		
Matches 702; Conservative	0;	Mismatches 50;	Indels 3;	Gaps 1;

OY	1	ATGAGGGGGCAGCCCACTGATTCATCTTCTGGCCACTTCCTCCTCGCCTTCTTCAATG	60
Dd	1478	ATGAGGGGGCAACCCCACTGATTCATCTTCTGGCCAATTTCCTCCTCGCATTTCTCAATG	1419
OY	61	GTTGTGCCCACTGTGCCCCGACACCCTGTACTGTCTTGACACCACCCCACTGCCA	120
Dd	1418	GTGTATTTCCCAGCTGTGCCCAAGCACCCCTGTGCCTTGTCTTGACACCACCCCACTGCCA	1359
OY	121	CAGGGGGTACCCCTGCTGCTGATGGCTGTGCTGTAAGTGTGTGCAGGAGGCTG	180
Dd	1358	CCGGGGGTACCCCTGCTGCTGATGGCTGTGCTGCTGAGTGTGTGCAGGAGGCTG	1299
OY	181	GGGAGTCTGCGAACCACTGCATGCTCGACCCCAAGCAGGGCCTGTTTGCAGCCT	240
Dd	1298	GGGAGTCTGCGAACCACTGCATGCTCGACCCCAAGCAGGGCCTGTTTGCAGCCT	1239
OY	241	GGGCGAGGCCCTGCGGCCCATGGGGCTGTGTCTTGTGATGAGATGACGGTAAGCTGT	300
Dd	1238	GGGCGAGGCCCACTGCGGCCCTGCTGTGTGCTTCTGAAGAGGATGACGGGAGCTGT	1179
OY	301	GAGGTGAATGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAGGTCCTG	360
Dd	1178	GAGGTGAATGGCCGAGGTACCTGGATGGGAGACCTTTAAACCCCAATTGCAGGTTTTG	1119
OY	361	TGCCGCTGTGATGACGGTGGCTTCACCTGCTGCCCTGTGCAGTGAGGATGTCCGGCTG	420
Dd	1118	TGCCGCTGTGATGACGGTGGTTTCACCTGCTGCCCTGTGCAGTGAGGATGTCCGGCTG	1059

QY	421	CCCAAGCTGGGACTGCCACCACGCCCCCACAAGAATATACAGGTGCCAGGAAGTCTGCCCCGAG	480
Dd	1058	CCCAGCTGGGACTGCCACCACGCCCCCACAAGAATAACAGGTGCCAGGAAGTCTGCCCCGAG	999
QY	481	TGGGTATGTGACCAGGAGTGTA--CACCGGCATCCAGCGCTCCACGGCGCAAGGACAC	537
Dd	998	TGGGTGTGTGACCAGGACAGTAGATGACGCGCGCAATCCAGCCCTCTCAGGCCCAAGGACAC	939
QY	538	CAACTTTCTGCCCTTGTCACTTCCTGCTGTGATGTCCTCTTGTCCAATTGGAGCACA	597
Dd	938	CAACTTTCTGCCCTTGTCACTTCCTGCAATGTGCCGATGGCCCCCTGTCCAACCTGGAAGCACA	879
QY	598	G CCTGGGGGCCCTGCTCAACCACTGTGGCGCTGGGCATAGCCACCCGAGTGTCCAACCAG	657
Dd	878	G CCTGGGGGCCCTGCTCAACCACTGTGGGTTGGGCATAGCCACCCGAGTGTCCAACCAG	819
QY	658	AACCGATTCTGCCA ACTGAGATCCAACGCGCGCTGTGTCTGCCACGACCCGTGCCTGGCA	717
Dd	818	AACCGATTCTGCCA ACTGAGATCCAACGCGCTGTGTCTGTCCAGACCCGTGCCTGGCA	759
QY	718	GCCAGGAGCCACAGCTCATGGAACAGTCTTTCTA	752
Dd	758	TCCAGGAGCCACGGCTCATGGAACAGTGCCTTCTA	724

RESULT 4	AF100778	1734 bp	mRNA	linear	ROD 17-DEC-1998
LOCUS	AF100778				
DEFINITION	Mus musculus connective tissue growth factor related protein WISP-2 (WISP2) mRNA, complete cds.				
ACCESSION	AF100778				
VERSION	AF100778.1	GI:4028578			

SOURCE	Mus musculus.
ORGANISM	Mus musculus

REFERENCE
AUTHORS
1 (bases 1 to 1734)
Pennica, D., Swanson, T. A., Welsh, J. W., Roy, M. A., Lawrence, D. A.,
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

TITLE

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)

MEDLINE	99061933
PUBMED	9843955
REFERENCE	2 (bases 1 to 1734)

JOURNAL Submitted (23-OCT-1998) **MOLECULAR Oncology, Genentech, Inc., 1 DNA**

FEATURES	Location/Qualifiers
source	1. .1734

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="C57MG"
/cell_type="epithelial"
/tissue_type="mammary"
/note="transformed by Wnt-1"
1.1734
/gene="Wisp2"
257.1012

```

```

/codon_start=1
/product="connective tissue growth factor related protein
WISP-2"
/protein_id="AAC96320.1"
/db_xref="GI:4028579"
/translation="MRGNPLHLAISFLCLISLMVYSQLCAPACPCPWTPEQCPGVP
LVLDGCGCCRYCARRLGESCDHLHVCDDPSQILVCQPGAGPSGRGAVCLFEEDDGSCVE

```


QY	301	GAGGTGAATGGCCGACAGTACTTGATGGAGAGACCCTTTAAACCCAATTGCAGGGTCCTG	360
Db	542	GAGGTGAACGGCCGACAGTACTTGATGGGAGAACCTTTAAACCAATTGCAGGGTTTTG	601
QY	361	TGCCCGTGTATGACGGGTGGCTTCACTGCGCTGCCGTGTGCAGTAGGATGTGCGGCTG	420
Db	602	TGCCCGTGTATGACGGGTGGTTTCACTGCGCTGCCGTGTGCAGTAGGATGTGCGGCTG	661
QY	421	CCCAGCTGGGACTGCCACGCCCCAGAAGAATACAGGTGCCAGGAAGTGTGCCCCGAG	480
Db	662	CCCAGCTGGGACTGCCACGCCCCAGAAGAATACAGGTGCCAGGAAGTGTGCCCCGAG	721
QY	481	TGGGTATGTGACCAGGGAGTGA---CACCGGCGATCCAGCGCTCCACGGCGCAGAGACAC	537
Db	722	TGGGTATGTGACCAGGCAGTGAATGCAGCCGCAATCCAGCCCTCTCAGCCCAAGACAC	781
QY	538	CAACTTTCTGCCCTTGTCACTOCTG CCTCTGCTGATGCTCCTTTGCCAATAITGGAGCACA	597
Db	782	CAACTTTCTGCCCTTGTCACTOCTG CATCTGCCGATGGCCCCCTGTCCAAACTGGAGCACA	841
QY	598	GCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAg	657
Db	842	GCTTGGGGCCCTGCTCAACCACTGTGGGTTGGGCATAGCCACCCGAGTATCCAACCAg	901
QY	658	AACCGATTCTGCCA ACTGGAGATCCAACGCCGCCCTGTGTCTGCCCAGACCCCTGCC TGGA	717
Db	902	AACCGATTCTGCCA ACTGGAGATCCAACGCCGCTGTGTCTGTCCAGACCCCTGCC TGGA	961
QY	718	GCCAGGAGCCACAGCTCATGGAACAGTGTCTTA 752	
Db	962	TCCAGGAGCCACGGCTCATGGAACAGTGTCTTA 996	

RESULT	6
LOCUS	AX076919
DEFINITION	Sequence 31 from Patent WO0105836.
ACCESSION	AX076919
VERSION	AX076919.1
KEYWORDS	GI:13121575
SOURCE	.
ORGANISM	human. Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1266) Botstein,D., Goddard,A., Gurney,A.L., Hillan,K.J., Roy,M.A. and Wood,W.I.
TITLE	Polypeptidic compositions and methods for the treatment of tumors
JOURNAL	Patent: WO 0105836-A 31 25-JAN-2001;
FEATURES	Genentech, Inc. (US)
source	Location/Qualifiers 1..1266 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	216 a 418 c 390 g 242 t
ORIGIN	

Query Match	67.8%;	Score 510.4;	DB 6;	Length 1266;
Best Local Similarity	79.9%;	Pred. No. 3.3e-114;		
Matches 601; Conservative	0;	Mismatches 151;	Indels 0;	Gaps 0;

QY	1	ATGAGGGGCAGCCCACTGATCATCTTCTG6CCACTTCCTTCTGTCTCATA TG	60
Db	10	ATGAGAGGCACACCGAAGACCCACTCTGCGCTTCCTCTCTGCTCTCTCAAAG	69
QY	61	GTCGTGCCCAGCTGTGCGGACACCCTGTACCTGTCTTGACACACCA CCAAGTGCCA	120
Db	70	GTGCGTACCCAGCTGTGCCCCGACACCACTGTA CCTGCCCCCTGGCCACCTCCC CGATGCCCG	129
QY	121	CAGGGGGGTACCCCTGTGCTGATGAGCTGTGGCTGCTGTAAAGTGTGTGACGGAGGCTG	180
Db	130	CTGGGAGTACCCCTGTGTCGTGATGCTGTGGCTGTCTGCCGGGTATGTGACA GCGCGGCTG	189

QY	181	GGGGAGTCTCGGACCACCTGCATGCTGCGACCCCGACGGCCCTGGTTGTGACGCT	240
Db	190	GGGGAGCCCTGCGACCAACTCCAGCTCTGCGACGCCAGCCAGGGCCCTGGTCTGCCAGCCC	249
QY	241	GGGGCAGGCCCTGGCGGCATGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT	300
Db	250	GGGGCAGGACCCGGTGGCCGGGGGGCCCTGTGCCCTCTTGGCAGAGGACGACAGCAGCTGT	309
QY	301	GAGGTGAATGCGCCGACGCTACCTGGATGGAGAGACCTTAAACCCAAATTCAGGGTCCTG	360
Db	310	GAGGTGAACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCACATCCGC	369
QY	361	TGCCGCTGTGATGACGGTGGCTTCACCTGCTGCCGCTGTGACGTGAGGATGTGCGGCTG	420
Db	370	TGCCGCTGCGAGGACGGCGGCTTCACCTGCTGCCGCTGTGACGCGAGATGTGCGGCTG	429
QY	421	CCGAGCTGGGACTGCCACGCCCCCAAGAAATACAGGTGCCAGGAAGTGTGCCCCGAG	480
Db	430	CCGAGCTGGGACTGCCACGCCCCCAAGAGGGTGCAGGTCTGGGCAAGTGTGCCCTGAG	489
QY	481	TGGGTATGTGACCAAGGAGTGTACACCGCGCATCCAGCGCTCCACGCGCGCAAGACACCAA	540
Db	490	TGGGTGTGCGGCCAAGGAGGGGAGCTGGGGACCCAGCCCTTCCAGCCCAAGGACCCCAAG	549
QY	541	CTTCTGCCCTGTGCACCTCCCTCTGCTGATGCTCTTGTCCAAATGGAGCACAGCC	600
Db	550	TTTCTGGCCCTGTCTCTTCCCTGCCCCCTGTGTCCCTGCCAGAAATGGAGCACAGGCC	609
QY	601	TGGGGCCCTGCTCAACCAACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAAC	660
Db	610	TGGGGACCCCTGCTCGACCAACCTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAGAAC	669
QY	661	CGATTCTGGCAACTGGAGATCCAACGCCGCCCTGTGTCTGCCACAGACCTTGCTGGCAGCC	720
Db	670	CGCTTCTGCCGACTGGAGACCCAGCGCCGCTGTGCTGTCCAGGCGCCCTGCCACCCCTCC	729
QY	721	AGGAGCCACAGCTCATGGAAACAGTGTCTTCTA	752
Db	730	AGGGGTGCGAGTCCACAAACAGTGTCTTCTA	761

RESULT 7					
AX464186					
LOCUS	AX464186	1266 bp	DNA	linear	PAT 16-JUL-2002
DEFINITION	Sequence 319 from Patent WO0140466.				
ACCESSION	AX464186				
VERSION	AX464186.1	GI:21899109.			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Baker, K.P., Beresini, M., Deforge, L., Desnoyers, L., Filvaroff, E., Gao, W.O., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.L., Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K., Wood, W.L. and Zhang, Z.				
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding same				
JOURNAL	Patent: WO 0140466-A 319 07-JUN-2001;				

Genentech Inc. (US)

FEATURES	Location/Qualifiers
source	1. .1266
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	216 a 418 c 390 g 24
ORIGIN	

Query Match	67.8%;	Score 510.4;	DB 6;	Length 1266;
· Best Local Similarity	79.9%;	Pred. No. 3.3e-114;		
Matches 601; Conservative	0;	Mismatches 151;	Indels 0;	Gaps 0;

QY 1 ATGAGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCCTCTGCCCTTCTCTCAATG 60

Db 10 ATGAGAGGACACACCGAAGACCACCTCTGCGCTTCTCCCTCTGCTCTCAAG 69
QY 61 GTGTGTGCCAGCTGTGCGGACACCTGTACTGTCTTGACACACCCAGTGCCCA 120
Db 70 GTGCGTACCAGCTGTGCGGACACACCTGTGCTGCGCTGCGCTGCGCTGCG 129
QY 121 CAGGGGTACCCCTGTGCTGTGATGGCTGTGCTGTAAAGTGTGTGACAGGAGCTG 180
Db 130 CTGGAGTACCCCTGTGCTGTGATGGCTGTGCTGCGGCTGCTGCGCTGCG 189
QY 181 GGGAGTCTGCGACACCTGCATGTCTGCGACCCAGCCAGGCGCTTGTGACGCT 240
Db 190 GGGAGCCCTGCGACCACTCCAGCTGTGCGACGCCAGCGGCTGTGCGACGCC 249
QY 241 GGGGACGCGCTGCGGCTGCGCTGTGCTGTCTTGGATGAGATGACGGTACTGT 300
Db 250 GGGGACGACCCGGTGGCGGGGGCGCTGTGCTTGGCAGAGGACGACAGCAGCTGT 309
QY 301 GAGGTGAATGCGCCGACGATGATGAGAGACCTTTAAACCAATTGACGGTCTG 360
Db 310 GAGGTGAACGCGCGCTGTATCGGAAGGGAGACCTTCCAGCCCACTGACATCCGC 369
QY 361 TGCCGCTGTGATGACGGTGTGCTTACCTGCTGCTGCTGCTGACATGAGATGCGGCTG 420
Db 370 TGCCGCTGTGAGGACGCGGCTTACCTGCGGTGCGGCTGTGACAGGAGATGCGGCTG 429
QY 421 CCCAGCTGGGACTGCCCCAGCGCCCAAGAAATACAGGTGCCAGAAAGTGTGCCCGAG 480
Db 430 CCCAGCTGGGACTGCCCCAGCGCGCTGAGAGGTGCTGCGGCAAGTGTGCCCTGAG 489
QY 481 TGGGTATGTGACCAAGGAGTGTACACCGCGATCCAGCGCTCCAGCGCGCAAGACCAA 540
Db 490 TGGGTGTGCGGCAAGGAGGAGGAGTGTGCGACCCAGCGCTTCCAGCGCAAGACCCAG 549
QY 541 CTTCCTGCGCTGTGCTACTCTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600
Db 550 TTTCTGTGCGCTGTGCTCTTCTTCCCTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCT 609
QY 601 TGGGGCCCTGCTCAACCACTGTGCGGTGGGCTAGGACACCGCGATGTCCAACAGAAC 660
Db 610 TGGGACCTGCTGACCACTGTGCGGTGGGCTAGGACACCGCGGTGTCCAACAGAAC 669
QY 661 CGATTCTGCCAAGTGTGAGATCCAAAGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 720
Db 670 CGCTTCTGCGACTGTGAGACCAAGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 729
QY 721 AGGAGCCACAGCTCATGGAACAGTGTCTTA 752
Db 730 AGGGGTGCGAGTCCACAACAGTGTCTTA 761

RESULT 8
AF083500 1283 bp mRNA linear PRI 04-NOV-1998
LOCUS Homo sapiens connective tissue growth factor-like protein
DEFINITION precursor, mRNA, complete cds.
ACCESSION AF083500
VERSION AF083500.1 GI:3462835
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1283)
AUTHORS Kumar,S., Hand,A.T., Connor,J.C., Dodds,R.A., Ryan,P.J.,
Trill,J.G., Fisher,S.M., Slemmon,J.R., Lipshutz,D.B.,
Bartholomew,V., James,I.E., Rleman,D.J., Gowen,M. and Lee,J.C.
IDENTIFICATION and cloning of CTGF-L from human osteoblasts, a
novel cysteine rich protein containing an IGF binding domain
JOURNAL Bone 23 (5), S240 (1998)
REFERENCE 2 (bases 1 to 1283)
AUTHORS Kumar,S.

TITLE Direct Submission
JOURNAL Submitted (11-AUG-1998) Bone & Cartilage Biology, UW 2109,
Smithkline Beecham, 709 Swedeland Road, King of Prussia, PA 19406,
USA
FEATURES
source Location/Qualifiers
1.1283
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q12-q13"
/cell_type="primary osteoblast"
9.761
/note="CTGF-L; encodes IGF binding (IGFB), Von Willebrand
factor type C (VMC) and thrombospondin type I (TSP1)
domains; member of the CCN (CTGF/cyrf61/nov) family; lacks
the fourth carboxy-terminal domain present in other
members of the CCN family"
/product="connective tissue growth factor-like protein
precursor"
/protein_id="AAC70350.1"
/db_xref="GI:3462836"
/translation="MRGTPKTHLLAFSLCLLSKVRQLCPPTCPWPPRPCLGVP
LVLDGCGCVRVCAARLGEPCDOLHVCASOGLVCPGAGPBGALCLLAEDSSGEY
NGRLYREGTFQPHCSIRKCEDEGFTCVPLCSVDYRLPSWDCPHPRVEVLKCPPE
WVCGGGLGTQPLPAQGPDSGLVSLPVGVPCEPNSWAWGPCSTTGLGMATRYSN
ONRFCLFQRLCLSRPDPSPGRSPNSAF"
BASE COUNT 235 a 418 c 389 g 241 t
ORIGIN
Query Match 67.8%; Score 510.4; DB 9; Length 1283;
Best Local Similarity 79.9%; Pred. No. 3.3e-114;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
QY 1 ATGAGGGGACCCACTGATCCATCTTGTGCGCACTTCTCTCTCTCTCTCAATG 60
Db 9 ATGAGAGGACACCGAAGACCACCTCTGCGCTTCTCCCTCTGCTCTCTCAAG 68
QY 61 GTGTGTGCCAGCTGTGCGGACACCTGTACTGTCTTGACACCAACCCAGTGCCCA 120
Db 69 GTGCGTACCAGCTGTGCGGACACCACTGTACTGTGCGCGCTGCGCGATGCCCG 128
QY 121 CAGGGGTACCCCTGTGCTGTGATGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
Db 129 CTGGAGTACCCCTGTGCTGTGATGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 188
QY 181 GGGAGTCTGCGACCACTGTGATGTGCGAACCCAGCGCGCTGCTGTGCAAGCT 240
Db 189 GGGAGCCCTGCGACCACTGTGCGAACCCAGCGCGCTGCTGTGCGCAAGCT 248
QY 241 GGGGACGCGCTGCGGCGCATGGGCTGTGCTGTGCTGTGATGAGATGACGGTACTGT 300
Db 249 GGGGACGACCGGCTGCGCGGGGCGCTGTGCTGTGCGAGAGGACAGACAGCAGCTGT 308
QY 301 GAGGTGAATGCGCGGACAGTACTGTGATGAGAGACCTTTAAACCAATTGACAGGTCTCTG 360
Db 309 GAGGTGAACGCGCGCTGTATCGGAAGGGAGAGACCTTCCAGCCCACTGACATCCGC 368
QY 361 TGCCGCTGTGATGACGGTGTGCTTACCTGCTGCGGCTGTGCAAGATGATGCGGCTG 420
Db 369 TGCCGCTGTGAGGACGCGGCTTACCTGTGCGGCTGTGCAAGATGATGCGGCTG 428
QY 421 CCCAGCTGGGACTGCCACAGCGCCCAAGAGAAATACAGGTGCCAGAAAGTGTGCCCGAG 480
Db 429 CCCAGCTGGGACTGCCACAGCGCGCTGAGAGGTGCTGTGCGAAGTGTGCGCTGAG 488
QY 481 TGGGTATGTGACCAAGGAGTGTACACCGCGATCCAGCGCTCCAGCGCGCAAGACCA 540
Db 489 TGGGTGTGCGGCAAGGAGGAGTGTGAGAACCAAGCGCGCTTCCAGCGCAAGACCCAG 548
QY 541 CTTCCTGCGCTGTGCTACTCTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600
Db 549 TTTCTGTGCGCTGTGCTCTTCCCTGCGCGCTGTGCTGTGCGCAAGATGAGACAGCGCC 608

QY	601	TGGGGCCCCCTGCTCAACCA	609	TGGGGACCCCTGCTCGACCA
QY	601	TGGGGCCCCCTGCTCAACCA	609	TGGGGACCCCTGCTCGACCA
Db	609	TGGGGACCCCTGCTCGACCA	668	TGGGGACCCCTGCTCGACCA
QY	661	CGATTCTGCGCAACTGGAGA	668	TGGGGACCCCTGCTCGACCA
Db	669	CGCTTCTGCGCAACTGGAGA	720	CGATTCTGCGCAACTGGAGA
QY	721	AGGAGCCACACAGTCATGGA	728	CGCTTCTGCGCAACTGGAGA
Db	729	AGGGGTCGCAGTCCACACAA	752	AGGAGCCACACAGTCATGGA

RESULT 9			
AR210322			
LOCUS	AR210322	1293 bp	DNA
DEFINITION	Sequence 13 from patent US 6387657.		linear
ACCESSION	AR210322		PAT 20-JUN-2002
VERSION	AR210322.1	GI:21512523	

ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1293)			
TITLE	Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Penicka,D., Roy,M.Ann. and Wood,W.I.			
JOURNAL	WISP polypeptides and nucleic acids encoding same			
FEATURES	Patent: US 6387657-A 13 14-MAY-2002;			
	Location/Qualifiers			
source	1..1293			
BASE COUNT	232 a 425 c 393 g 243 t			
ORIGIN	/organism="unknown"			

Query Match	67.88;	Score 510.4;	DB 6;	Length 1293;
Best Local Similarity	79.98;	Pred. No. 3.3e-114;		
Matches 601; Conservative	0;	Mismatches 151;	Indels 0;	Gaps 0;

QY	I	ATGAGGGGCACGCCACTGATCCATCTTCTGCGCACCCTTCCTCTGTGCTTCTCAATG	60
Db	22	ATGAGAGGCACACCGAAGACCACCTCTGGCCCTTCTCCCTCTGTGCTCTCAAAAG	81
QY	61	GTTGTGTGCCAGCTGTGCGCGACCCCTGTACCTGTCTTGACACACCACCGTGCCCA	120
Db	82	GTCGTAACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACTCCCAGATGCCG	141
QY	121	CAGGGGTACCCCTGTGCTGTGATGGCTGTGGCTGCTGTAAAGTGTGTACACGGAGCTG	180
Db	142	CTGGAGTAGCCCCCTGTGCTGTGATGGCTGTGGCTGCTGCCGGGTATGTGCACGGCGCTG	201
QY	181	GGGGAGTCTTGCACCACTGCATGTCTGCGACCCACGCCAGGGCCCTGTGTCTCAGCCT	240
Db	202	GGGGAGCCCTTGCACCACTGCACGTCTGCGACGCCAGCCAGGGCCCTGTGTGCCAGCCC	261
QY	241	GGGGCAGGCCCTGGCGCGCCATGGGGCTGTGTCTCTTGATGAGATGACGGTAGCTGT	300
Db	262	GGGGCAGGACCCGGTGGCGCGGGGCCCTGTGTCTCTTGAGAGAGACGACGACGACTGT	321
QY	301	GAGGTGAATGGCCGACAGTACCTGAGATGGAGAGACCTTTAAACCAATTGCAGGGTCCCTG	360
Db	322	GAGGTGAACGGCCGCTGTATCGGGAAGGGAGAGACCTTCCAGCCCACTGCAGCATCCGC	381
QY	361	TGCCGCTGTGATGACGGTGGCTTCACTTGCCCTGCCGCTGTGCAGTGAAGATGTGGCGCTG	420
Db	382	TGCCGCTGCGAGGACGGCGGCTTCACTTGCGTGGCGCTGTGCAGCAGGATGTGGCGCTG	441
QY	421	CCCAGCTGGGACTGCCACACGCCCCCAAGAGAATACAGGTGCCAGGAAAGTCTGCCCCGAG	480
Db	442	CCCAGCTGGGACTGCCCCACCCCAAGAGGGTCAAGTCTTGCGCAAGTGTGCCCTGAG	501
QY	481	TGGGTATGTGACCAGGAGATGACACCGCGATCCAGCGCTCCAGCGGCAAGGACACCAA	540

Db 502 TGGGTTGCGGCCAAGGAGGGGGACTGGGGACCCAGCCCCCTTCCAGCCCAAGGACCCCCAG 561

QY 541 CTTTCTGCCCTTGTCACTCCCTGCCTCTGCTGATGCTCCTGTGCCAATTTGGACACAGCC 600
||||| ||||||| || | ||||| || ||||| ||||| |||

Db 562 TTTTCTGGCCCTGTCTCTTCCCTGCCCCCTGGTGTCCCTCGCCAGATGGAGCACGGCC 621

QY 601 TGGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAAC 660
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 622 TGGGGACCCCTGCTCGACCACCTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAGAAC 681

QY 661 CGATTCTGCCAACTGGAGATCCAACGCCGCCCTGTGTCTGCCCCAGACCCCTGGCAGCC 720
|| ||||||| ||||||| ||| ||||||| ||| ||||| ||||||| | ||

Db 682 CGCTTCTGCGCACTGGAGACCACAGCGCGCCCTGTGCTGTCCAGGCGCCCTGCCACCCTCC 741

QY 721 AGGAGCCACAGCTCATGGAAACAGTGCTTTCTA 752
||| | | ||| || ||||||| |||||

Db 742 AGGGGTCGCACTCCACAAACAGTGCCCTTCTA 773

RESULT 10			
AR210323/c			
LOCUS	AR210323	1293 bp	DNA
DEFINITION	Sequence 14	from patent US 6387657.	linear
ACCESSION	AR210323		
VERSION	AR210323.1	GI:21512524	PAT 20-JUN-2002

ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1293)
TITLE	Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
JOURNAL	Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.
FEATURES	WISP polypeptides and nucleic acids encoding same
Source	Patent: US 6387657-A 14 14-MAY-2002;
	Location/Qualifiers
	1. .1293
BASE COUNT	/organism="unknown"
ORIGIN	243 a 393 c 425 g 232 t

Query Match	67.8%;	Score 510.4;	DB 6;	Length 1293;
Best Local Similarity	79.9%;	Pred. No. 3.3e-114;		
Matches 601;	Conservative 0;	Mismatches 151;	Indels 0;	Gaps 0;
QY 1	ATGAGGGGAGCCCACTGATTCATCTTCTGCGCACCTTCCTCTGCTCTCTCAATG	60		
Db 1272	ATGAGAGGCACACCGAAGACCACCTCTGCGCTTCCTCTCTGCTCTCTCAAG	1213		
QY 61	GTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTGGACACACCCAGTGCCCA	120		
Db 1212	GTGCGTACCCAGCTGTGCCCGACACCATGTACCTGTGCCCTTGCGCACCTCCCGATGCCG	1153		
QY 121	CAGGGGCTACCCCTGTGTCTGATGGCTGTGGCTGTCTTAAGTGTGTGCACGAGGCTG	180		
Db 1152	CTGGGAGTACCCCTGTGTCTGATGGCTGTGGCTGTGTGCCGGGTATGTGCACGGCGGCTG	1093		
QY 181	GGGAGTCTTGGCAGCACCTGCATGCTGCGACCCCCCAGGCGCTGTTGTACGCT	240		
Db 1092	GGGGAGCCCTGGCAGCACTCCACGCTCTGCGACGCCAGGGCGCTGTTGTGCCAGCCC	1033		
QY 241	GGGGCAGGCCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT	300		
Db 1032	GGGGCAGGACCCGTTGGCGGGGGGGCCCTGTGCTCTTGGCAGAGGACGACGACGCTGT	973		
QY 301	GAGTGAATGGCCGACGTAACCTGGATGGAGAGACCTTAAACCAATTGCAGGCTCTG	360		
Db 972	GAGTGAACGGCCGCTGTATCGGGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGC	913		
QY 361	TGCCGCTGTGATGACGGTGGCTTCACCTGCTGCCCTGTGTGCAGTGAGGATGTGCCGCTG	420		
Db 912	TGCCGCTGCGAGGACGGCGGCTTCACCTGCTGCCCTGTGTGCAGCGAGGATGTGCCGCTG	853		
QY 421	CCCAGCTGGGACTGCCACGCCCCCAAGAGATACAGGTGCCAGGAAGTGTGCCCCGAG	480		

|||||
Db 852 CCCAGCTGGAGCTGCCCCCACCAGAGGGTGGAGTCTGTGGCAAGTGTGCTGAG 793
Qy 481 TGGGTATGTGACACGAGGAGTGACACCGCGATCCAGCGCTCCACGGCGCAAGACACCAA 540
Db 792 TGGGTGTGGCGCAAGAGGAGGGAGCTGGGAGCCAGCCGCTCCAGGCCCAAGACCCAG 733
Qy 541 CTTTCTGCGCTGTGACTCTGCTGCTGTGATGCTCTGTCTCAATGTGAGCAGAGCC 600
Db 732 TTTTCTGGCTGTGCTCTCTCCCTGCCCCCTGTGCTCCAGATGAGCAGAGCC 673
Qy 601 TGGGGCCCTGCTCAACACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACAGAAC 660
Db 672 TGGGAGCCCTGCTCGACACCTGTGGGCTGGGCATGCCCCACCCGGGTGTCCAACAGAAC 613
Qy 661 CGATTCTGCCAAGTGGAGATCCAAACCGCGCTGTGTCTGCGCAGACCTGCGTGGCAGCC 720
Db 612 CGCTTCTGCGACTGGAGACCCAGCGCGCTGTGCTGTCCAGGCGCTGCCCCACCTCC 553
Qy 721 AGGAGCCACAGCTCATGGAACAGTGTCTTA 752
Db 552 AGGGTGGCAGTCCACAAACAGTGTCTTA 521

RESULT 11
AF074604 1309 bp mRNA linear PRI 18-JUL-1998
LOCUS Homo sapiens connective tissue growth factor-related protein
DEFINITION precursor (CT58) mRNA, complete cds.
ACCESSION AF074604
VERSION AF074604.1 GI:3328191
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1309)
AUTHORS Rowles,J. and Gendler,S.
TITLE CT58, a new member of the connective tissue growth factor family,
interacts with the breast cancer associated mucin MUC1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1309)
AUTHORS Rowles,J. and Gendler,S.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1998) Biochemistry and Molecular Biology, Mayo
Clinic Scottsdale, 13400 E. Shea Blvd., Scottsdale, AZ 85259, USA
FEATURES
source
1. 1309
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Hela"
1. 1309
/gene="CT58"
7. 759
/note="CT58"
/note="contains three of four modules found in growth
regulators related to connective tissue growth factor as
described in FEBS letters.327:125-130,1993; identified in
yeast two-hybrid screen using the epithelial mucin MUC1 as
bait"
/codon_start=1
/product="connective tissue growth factor-related protein
precursor"
/protein_id="AAC26794.1"
/db_xref="GI:3328192"
/translation="MRGTPKTHLLAFLSLCLSLVRLQICPTPTCPWPPRPCLGVP
LVLDGCGCCRVCAARLGEPCDQLHVCDAISGLVGPAGPGGAGALCLLAEDDSCEV
NGRLVREGETFQPHCSIRKCEDGGFTCVPLCSLDVRLPSWDCPPHRRVEVLGKCCPE
WVCGGGIGTQPLPAQGPDSGLVSLPVPVCPENSTAWGPCSTTCGLMATRVSN
ONRFRLETFORRLCLSRPCPPSRGRSPONSAP"

BASE COUNT 261 a 418 c 387 g 242 t 1 others

Query Match 67.8%; Score 510.4; DB 9; Length 1309;
Best Local Similarity 79.9%; Pred. No. 3.3e-114;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

Qy 1 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTGCTGCTCTCAATG 60
Db 7 ATGAGAGGACACCGAAGACCACCTCTGCGCTTCTCTCTCTGCTCTCTCAAG 66
Qy 61 GTGTGCGCAGCTGTGCGGAGACACCTGTACCTGTCTCTGAGACACCCAGTGCCCA 120
Db 67 GTGCTACCCAGCTGTGCGGAGACACCTGTACCTGTGCGGAGACACCCAGTGCCCA 126
Qy 121 CAGGGGTACCCCTGTGCTGTGATGGCTGTGGCTGTGTAAGTGTGTGACAGAGGCTG 180
Db 127 CTGGAGTACCCCTGTGCTGTGATGGCTGTGGCTGTGTAAGTGTGTGACAGAGGCTG 186
Qy 181 GGGAGTCTGCGACACCTGCATGTCTGCGACCCAGCCAGGCGCTGTGTGACGCT 240
Db 187 GGGAGCCCTGCGACACCTGCATGTCTGCGACCCAGCCAGGCGCTGTGTGACGCT 246
Qy 241 GGGGAGCCCTGCGGCGCATGGGCTGTGTCTCTGTGATGAGATGACGCTGTGT 300
Db 247 GGGGAGGACCCGGTGGCGCGGCGGCGGCTGTGTCTGTGAGAGGACGACAGCAGCTGT 306
Qy 301 GAGGTGAATGCGCCGAGGATACCTGTGAGAGAGACCTTTAAACCAATTGCAGGCTCTG 360
Db 307 GAGGTGAACGCGCCGCTGTATCGGAGAGGAGAGACCTTCACGCCCACTGCAGCATCCGC 366
Qy 361 TGGCGCTGTGATGACGGGTGGCTTTCACCTGCGCTGCGCTGTGAGAGATGCGGCTG 420
Db 367 TGGCGCTGTGATGACGGGTGGCTTTCACCTGCGCTGCGCTGTGAGAGATGCGGCTG 426
Qy 421 CCCAGCTGGAGCTGCCACGCGCCCAAGAGATACAGGTGCCAGGAAGTGTGCCCGAG 480
Db 427 CCCAGCTGGAGCTGCCACGCGCCCAAGAGAGGTGCCAGGTGCCAGGTGTGCCCTGAG 486
Qy 481 TGGGTATGTGACACGAGGAGTGACACCGCGCATCCAGCGCTCCAGCGCGCAAGACACCA 540
Db 487 TGGGTGTGGCGCAAGAGGAGGGAGCTGGGAGACCCAGCCCTTCCAGCCCAAGAGACCCAG 546
Qy 541 CTTTCTGCGCTGTGACTCTGCTGTCTGTGATGCTCTCTTGTCCAAATTGGAGCAGACGC 600
Db 547 TTTTCTGGCTGTGCTCTTCCCTGCGCGCTGTGTGCTCCCTGTCCAGATGAGCAGCGCC 606
Qy 601 TGGGGCCCTGCTCAACACCTGTGGCTGGGCATAGCCACCCGAGTGTCCAACAGAGAC 660
Db 607 TGGGAGCCCTGCTCGACACCTGTGGCTGGGCATGGCCACCCGGGTGTCCAACAGAGAC 666
Qy 661 CGATTCTGCCAAGTGGAGATCCAAACCGCGCTGTGTCTGCGCAGACCTGCGTGGCAGCC 720
Db 667 CGCTTCTGCGACTGGAGACCCAGCGCGCTGTGCTGTGCTGTGCGCAGCCCTGCCACCCCTCC 726
Qy 721 AGGAGCCACAGCTCATGGAACAGTGTCTTA 752
Db 727 AGGGTGGCAGTCCACAAACAGTGTCTTA 758

RESULT 12
AF100780 1427 bp mRNA linear PRI 17-DEC-1998
LOCUS Homo sapiens connective tissue growth factor related protein WISP-2
DEFINITION (WISP2) mRNA, complete cds.
ACCESSION AF100780
VERSION AF100780.1 GI:4028582
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1427)
AUTHORS Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A.,
Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C.,
Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D.,

TITLE Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.
WISP genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)
MEDLINE 99061933
PUBMED 9843955
REFERENCE 2 (bases 1 to 1427)
AUTHORS Penlica,D.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA
FEATURES
source Location/Qualifiers
1..1427
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q12-q13.1"
/tissue_type="lung"
/dev_stage="fetus"
1..1427
/gene="WISP2"
156..908
/gene="WISP2"
/codon_start=1
/product="connective tissue growth factor related protein WISP-2"
/protein_id="AAC96322.1"
/db_xref="GI:4028583"
/translation="MRGTPKTHLAFSLCLISKVKTQLCPPTCPMPPPRCPLGVPLVDGCGCCRVCAARRLGEPCDQLHVCASQGLVCOGAGPGRGALCLLAEDSSCEVNGRLYREGTEFQPHCSIRCRCEGDGFTCVLPSEDEVRLPSWDCPHPRRVEYVSKCPEWVGQGGGLGTQPLPAQGPQFSGLVSLPPGVPCPEWSTAWGPCSTTCGLMATRVSNQNRFCRLTQRLCLSRPCPPSRGRSPONSAP"

BASE COUNT 261 a 468 c 417 g 280 t 1 others
ORIGIN

Query Match 67.8%; Score 510.4; DB 9; length 1427;
Best Local Similarity 79.9%; Pred. No. 3.2e-114;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGCGAGCCCACTGATTCATCTTCTGGCCACTTCCTTCTGCTCTCAATG 60
Db 156 ATGAGAGGGCACACCGAAGACCCTCTGCGCTTCTCTCTCTGCTCTCAAG 215
QY 61 GTGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGAGACACCCAGTGCCA 120
Db 216 GTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCG 275
QY 121 CAGGGGGTACCCCTGTGTGTGATGGCTGTGGCTGTGTAAGTGTGTGACGAGCGTG 180
Db 276 CTGGGAGTACCCCTGTGTGTGTGATGGCTGTGGCTGTGCGGGGTATGTGACGCGGCGTG 335
QY 181 GGGGAGTCTTGCGACCACTGCATGTCTGCGACCCCGACGCGCCCTGTTGTCAAGCT 240
Db 336 GGGGAGCCCTGCGACCACTCCAGTCTGCGACGCCAGCGGCTGTCTGCCAGCCC 395
QY 241 GGGGAGGCCCCCTGGCGGCCATGGGGCTGTGTCTCTTGATGAGGATGACGGTAGCTGT 300
Db 396 GGGGAGGAGCCCGGTGGCGGGGGGCCCTGTGCTCTTGCGAGAGGACGACGACGACGCTGT 455
QY 301 GAGGTGATGGCCGCGAGGATACCTGGATGGAGAGACCTTTAAACCAATTGACGGGCTCTG 360
Db 456 GAGGTGAACGGCGCGCTGTATCGGGAAAGGGAGACCTTCAGCCCTGACGACATCCGC 515
QY 361 TGGCGCTGTATGACGGGTGCTTCACTGCGCTGCGGCTGTGACGTGAGGATGTGGCGCTG 420
Db 516 TGGCGCTGCGAGGAGCGGGCTTCACTGCGCTGCGGCTGTGACGAGGATGTGGCGCTG 575
QY 421 CCCAGCTGGAGTGGCCACGCCCCAAGAGATACAGGTGCGCAGGAAAGTGCTGCCCGAG 480
Db 576 CCCAGCTGGAGTGGCCACGCCCCAAGAGAGGAGGTGAGAGGTCTCTGGGCAAGTGTGCCCTGAG 635

QY 481 TGGGTATGTGACGAGGAGTGACACCGCGGATCCAGCGCTCCACGGCGCAAGACCA 540
Db 636 TGGGTGTGCGGCCAAGAGGGGAGCTGGGGACCCAGCCCTTCACGCCCAAGACCCAG 695
QY 541 CTTTCTGCCCTTGTACATCTGCTCTGCTGATGTCTCTGTCCAATTTGAGACAGACC 600
Db 696 TTTTCTGGCCTGTCTCTCCCTGCCCCCTGTGTCCCTGCCAGATGAGACAGGCC 755
QY 601 TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACGAGAC 660
Db 756 TGGGGAACCTGTCTGCACCACTGTGGGCTGGGCATGAGCCACCCGGGTGTCCAACGAGAC 815
QY 661 CGATTCTGCCAAGTGGAGATCCACACGCGGCTGTGTCTGCCCCAGACCCCTGGCAGCC 720
Db 816 CGCTTCTGCCGACTGGAGACCCAGCGCGGCTGTGCTGTCCAGGCGCTGCCACCTCC 875
QY 721 AGGAGCCACAGCTCATGGAACAGTGCTTTCTA 752
Db 876 AGGGGTGCGAGTCCACAAACAGTGCTTTCTA 907

RESULT 13
BC017782
LOCUS
DEFINITION Homo sapiens, WNT1 inducible signaling pathway protein 2, clone
ACCESSION MGC:22271 IMAGE:4691574, mRNA, complete cds.
VERSION BC017782
KEYWORDS BC017782.1 GI:17389482
SOURCE MGC.
ORGANISM Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1450)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdepaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAL Plate: 36 Row: m Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507922.

FEATURES
source Location/Qualifiers
1..1450

/organism="Homo sapiens"
/db_xref="LocusID:8839"
/db_xref="taxon:9606"
/clone="MGC:22271 IMAGE:4691574"
/tissue_type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
165..917
/codon_start=1
/product="WNT1 inducible signaling pathway protein 2"
/protein_id="AAH17782.1"
/db_xref="GI:17389483"

BASE COUNT	272 a	457 c	441 g	280 t
ORIGIN	/translation="MRGTPKTHLAFSLCLLQSKVRTQLCPTPCCTCPMPPPRCPLGVP LVLDGCGCCRYCARLIGEPQDLHYCDIAQWQCPAGPGRGALCLIAEDSSCEV NGRLYREGTEFQPHCSIRRCEDGEGFTVPLQSDVRLRPSWDCPHPRVEVIGKCCPE WVGCGGGGLGTQPLPAQGPQFSGLVSLPPGVPCPEWSTAWGCPSTCGLMATRVSN QNRFCRLETOGRLCLSRPCPPSRGRSPNSAF"			

Query Match	67.8%;	Score 510.4;	DB 9;	Length 1450;
Best Local Similarity	79.9%;	.Pred. No. 3.2e-114;		
Matches 601; Conservative	0;	Mismatches 151;	Indels 0;	Gaps 0;

QY	1	ATGAGGGGCGACGCCACTGATCCATCTTCTGGCCACTTCGCTTCTCTCAATG	60
Db	165	ATGAGAGGCGACACCGAAGACCACCTCTCGGECTTCTCCGCTCTGCTCTCAAG	224
QY	61	GTCGTGCCCCAGCTGTGCGCGACACCCCTGTAACCTGTCTTGGACACACCAGTGC	120
Db	225	GTCGTAACCCAGCTGTGCGCGACACCATGTACCTGCCCCCTGGCCACACTCCCGATGCCG	284
QY	121	CAGGGGCTACCCCTGCTGTGCTGATGCGCTGTGGCTGCTGTAAAGTGTGTGACGAGGCTG	180
Db	285	CTGGGAGTACCCCTGTGCTGTGATGCGCTGTGGCTGCTGCGGGTATGTGACGCGGCTG	344
QY	181	GGGGAGTCTCGGACACCTGCATCTCTGCGAACCCCGAGCGGCGCTGTGTTCAGCCT	240
Db	345	GGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCGAGGGGCTGTCTGCCAGCC	404
QY	241	GGGGCAGGCCCCCTGGCGGCCATGGGCGTGTGTCTCTTGGATGAGATGACGGTAGCTGT	300
Db	405	GGGGCAGGACCCCGGTGGCGCGGGGCGCTGTGCTCTTGGCAGAGACGACAGCAGCTGT	464
QY	301	GAGGTGAATGGCCGCGCAGGTACCTGGATGGAGAGACCTTTAACCAATTGACAGGTCCTG	360
Db	465	GAGGTGAACGGCGCGCTGTATCGGGAAGGGAGACCTTCAGCCCCACTGCAGCATCCGC	524
QY	361	TGCCGCTGTGATGACGCTGGCTTCACTGCTGCGCGCTGTGACGTGAGATGTGCGGCTG	420
Db	525	TGCCGCTGCGAGGACGGCGCTTCACTGCTGCGCTGTGACGAGATGTGCGGCTG	584
QY	421	CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAGGAAGTCTGCCCCGAG	480
Db	585	CCCAGCTGGGACTGCCACCCACCCAGAGAGGGTCGAGGTCCGTGGGCAAGTCTGCCCCTGAG	644
QY	481	TGGGTATGTGACCAAGGAGTGCACCGCGCATCCAGCGCTCCACGCGCGCAAGGACACCAA	540
Db	645	TGGGTGTGCGGCCAAGGAGGGGAGCTGGGGAACCCAGCCCCCTTCACAGCCCAAGGACCC	704
QY	541	CTTCTGCCCCCTGTGCACCTCCGCTCTGCTGATGCTCTCTGTCCAAATGGAGCACAGCC	600
Db	705	TTTCTGCGCTTGTCTTCCCTGCCCCCTGCTGTCCCTGCCAGAAATGAGCACGCGCC	764
QY	601	TGGGGCCCCCTGCTCAACCCACCTGCGGCTGGGCGATAGCCAGCCGAGTGTCCAACCAAGAC	660
Db	765	TGGGGACCCCTGCTGCACCCACCTGTGGGCTGGGCGATGGCCAGCCGGGTGTCCAACCAAGAC	824
QY	661	CGATTCTGCCAAGCTGAGATCCAAAGCCGCGCTGTGTCTGCCGACAGACCTTGGCAGCC	720
Db	825	CGCTTCTGCCGACTGGAGACCCAGCGCGCTGTGCTGTGACAGGCGCTGTGCCACCCCTCC	884
QY	721	AGGAGCCACAGCTCATGGAAACAGTCTTCTTA	752
Db	885	AGGGGTCCAGTCCACAAACAGTCCCTTCTTA	916

RESULT 14			
AR210337			
LOCUS	AR210337	738 bp	DNA
DEFINITION	Sequence 38 from patent us 6387657.		linear
ACCESSION	AR210337		PAT 20-JUN-2002
VERSION	AR210337.1	GI:21512542	
KEYWORDS			
SOURCE	Unknown.		

ORGANISM	Unknown.
	Unclassified.
REFERENCE	1 (bases 1 to 738)
AUTHORS	Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.
TITLE	WISP polypeptides and nucleic acids encoding same
JOURNAL	Patent: US 6387657-A 38 14-MAY-2002;
FEATURES	Location/Qualifiers
source	1..738
	/organism="unknown"
BASE COUNT	104 a 272 c 238 g 124 t
ORIGIN	

Query Match	66.6%;	Score 501.2;	DB 6;	Length 738;
Best Local Similarity	79.9%;	Pred. No. 6.2e-112;		
Matches 590; Conservative	0; Mismatches 148;	Indels 0;	Gaps 0;	

QY	13	CCACTGATCCATCTTTGCGCACTTCCCTCTGCTGCTCTCTCAATGGTGTGTCGCCAG	72
Db	1	CGAAGACCCACCTCTGGCCCTTCTCCCTCTCTGCTCTCTCAAGGTGCGTACCAG	60
QY	73	CTGTGCCGGACACCCCTGTACCTGTCTTGACACACCACCAGTGCACAGGGGTACC	132
Db	61	CTGTGCCGACACCATGTACCTGTCCCTTGCCACCTCCCGATGCCGCTGGAGTACC	120
QY	133	CTGTGCTGGATGGCTGTGGCTGTAAAGTGTGTGCACGAGCGTGGGGAGTCTGC	192
Db	121	CTGTGCTGGATGGCTGTGGCTGTCCGGTATGTGCACGCGCGCTGGGGAGCCCTGC	180
QY	193	GACCACCTGCATGTCTGCAGCCCAAGCCAGGGCCCTGGTTGTACGCTGGGGCAGGCCCT	252
Db	181	GACCAACTCCACGTCTGCCAGCCCAAGCCAGGGCCCTGGTCTGCCAGCCCGGGCAGGACCC	240
QY	253	GGCGGCCATGGGCTGTGTCTCTTGGATGAGGATGACCGTACCTGTGAGGTGAATGGC	312
Db	241	GGTGGCCGGGGGGCCCTGTGCTCTTGGCAGAGGACGACACGACTGTGAGGTGAACGGC	300
QY	313	CGCAGGTACCTGGATGAGAGACCTTAAACCCAAATTGCAGGGTCTGTGCCGCTGTGAT	372
Db	301	CGCTGTATCGGGAAGGGGAGACCTTCCAGCCCACTGCAGCATCTGCTGCCGTGCGAG	360
QY	373	GACGGTGGCTTACCTGCTGCCGCTGTGCAGTGAGGATGTGGGCTGCCAGCTGGGAC	432
Db	361	GACGGCGGCTTACCTGCTGCCGCTGTGCAGCGAGGATGTGGGCTGCCAGCTGGGAC	420
QY	433	TGCCCAGCCCCAAGAAATACAGGTGCCAGGAAGTGTGCCCGAGTGGGTATGTGAC	492
Db	421	TGCCCCCAACCCAGAGGGTCAAGTCTTGGCAAGTGTGCTGCCCTGAGTGGGTGTGCGGC	480
QY	493	CAGGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGAGACACCACTTCTGCCCTT	552
Db	481	CAAGGAGGGGAGCTGGGGACCCACCCCTTCCAGCCCAAGGACCCCACTTCTTGCCCTT	540
QY	553	GTCACCTCCTGCTCTGCTGATGCTCTTGTGCCAAATTGGAGCACAGCTGGGGCCCTGC	612
Db	541	GTCCTCTTCCCTGCCCCCTGTGTCTCCCTGCCCAAGATGAGCACGCGCTGGGGACCCCTGC	600
QY	613	TCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAACCGATTCTGCCAA	672
Db	601	TGACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAACCGCTTCTGCCGA	660
QY	673	CTGAGATCCAACGCGCCTGTCTTGCCAGACCCCTGCTGGCAGGAGGACCAAGC	732
Db	661	CTGAGAGCCACGCGCCCTGTGCTGTCCAAGGCCCTGCCACCCCTCAGGGGTGCGAGT	720
QY	733	TCATGGAACAGTGCCTTC	750
Db	721	CCACAAACAGTGCCTTC	738

RESULT 15	
AR210338	
LOCUS	AR210338
	841 bp
	DNA
	linear
	PAT 20-JUN-2002

DEFINITION Sequence 39 from patent US 6387657.

ACCESSION AR210338

VERSION AR210338.1 GI:21512543

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 841)

AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,

Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.

TITLE WISP polypeptides and nucleic acids encoding same

JOURNAL Patent: US 6387657-A 39 14-MAY-2002;

FEATURES Location/Qualifiers

1..841

source /organism="unknown"

BASE COUNT 124 a 297 c 280 g 140 t

ORIGIN

Query Match 65.9%; Score 496.4; DB 6; Length 841;

Best Local Similarity 79.7%; Pred.No.9.le-111;

Matches 599; Conservative 0; Mismatches 151; Indels 2; Gaps 1;

QY 1 ATGAGGGGAGCCGACCTGATCATCTTCTGGCCACTTCCTCCTCTGCTTCTCAATG 60

Db 12 ATGAGAGGCACACCGAAGACCCTCTGCGCTTCTCCCTCTGCTCTCTCAAG 71

QY 61 GTCTGTGCCAGCTGTGCCGACACCTGTACCTGTCTCTTGACACACCCAGTGCCCA 120

Db 72 GTGGCTACCCAGCTGTGCCGACACCATGTACCTGTCCCTGGCCACCTCCCGATGCCG 131

QY 121 CAGGGGGTACCCCTGTGCTGATGGCTGTGGCTGCTGTAAGTGTGTGCACGGAGGCTG 180

Db 132 CTGGGAGTACCCCTGTGCTGATGGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 191

QY 181 GGGGAGTCTTCGACCACTGATGTCGACACCCAGCCAGGGCTGTTGTACAGCT 240

Db 192 GGGGAGCCCTGCGACCACTGCTGCGACGCGCAGGGCTGCTGCTGCTGCTGCTGCTGCTG 251

QY 241 GGGGAGGGCCCTGGCGGCACTGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 300

Db 252 GGGGAGGAGCCCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 311

QY 301 GAGGTGAATGGCCGAGGTACCTGATGAGAGACCTTTAAACCAATTGCAGGGTCTCTG 360

Db 312 GAGGTGAACGGCCGCTGTATGAGGAGAGGGAGACCTTCCAGCCCTGACAGCATCCGC 371

QY 361 TGCCGCTGTGATGACGCTGCTTCACTGCTGCCGCTGTGAGTGAAGATGTGCGGCTG 420

Db 372 TGCCGCTGCGAGGAGCGGCGGCTTCACTGCTGCCGCTGTGAGTGAAGATGTGCGGCTG 431

QY 421 CCCAGCTGGGACTGCCACCGCCCAAGAGATACAGGTGCCAGGAAAGTCTGCCCGAG 480

Db 432 CCCAGCTGGGACTGCCACCGCCCAAGAGGAGGTGAGGTCTGGGCAAGTCTGCCCTGAG 491

QY 481 TGGGTATGTGACGAGGAGTGAACGGCGATCCAGCGCTCCACGGCGCAAGGACACCAA 540

Db 492 TGGGTATGTGACGAGGAGTGAACGGCGATCCAGCGCTCCACGGCGCAAGGACACCAA 549

QY 541 CTTTCTGCCCTTGTCACTCTCTGCTGATGCTCTTGTCAATTTGAGACAGCC 600

Db 550 TTTTCTGCCCTTGTCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609

QY 601 TGGGGCCCTGTCTAACACCTTGGGCTGGGATAGCCACCCAGTGTCCACACGAAAC 660

Db 610 TGGGGACCTGTCTGACACCTTGGGCTGGGATAGCCACCCAGTGTCCACACGAAAC 669

QY 661 CGATTCTGCCAAGTGAAGATCAACGGCGGCTGTGTGCTGCCAGACCCCTGCGCAGCC 720

Db 670 CGATTCTGCCAAGTGAAGATCAACGGCGGCTGTGTGCTGCCAGACCCCTGCGCAGCC 729

QY 721 AGGAGCCACAGCTCATGGAACAGTGTCTTA 752

Db 730 AGGGGTGCGAGTCCACAAACAGTGTCTTA 761

Search completed: July 29, 2003, 00:48:38
Job time : 2093.77 secs

XX
PS Claim 5; Fig 1; 108pp; English.

CC The invention provides a rat heparin-induced GCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents the coding sequence of rat HICP.

sq sequence 753 BP; 132 A; 235 C; 230 G; 156 T; 0 other;

Query Match	100.0%;	Score 753;	DB 20;	Length 753;
Best Local Similarity	100.0%;	Pred. No. 2.5e-193;		
Matches 753; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	ATGAGGGGCGACGCCACTGATCCATCTTCTGGCCACTTGGCTTCCCTCTGCTCTCAATG	60
Db	1	ATGAGGGGCGACGCCACTGATCCATCTTCTGGCCACTTGGCTTCCCTCTGCTCTCAATG	60
QY	61	GTCGTGCCAGCTGTGCCGACACCTGTACCTGTCTGTGGACACACCCAGTGCCCA	120
Db	61	GTCGTGCCAGCTGTGCCGACACCTGTACCTGTCTGTGGACACACCCAGTGCCCA	120
QY	121	CAGGGGGTACCCCTGTGTGATGGATGGCTGTGGCTGTGTAAAGTGTGTGACGAGGCTG	180
Db	121	CAGGGGGTACCCCTGTGTGATGGATGGCTGTGGCTGTGTAAAGTGTGTGACGAGGCTG	180
QY	181	GGGAGTCTCTGGACACCTGCATGTCGCCAGCCCCAGGCGGCTGGTTGTCAAGCT	240
Db	181	GGGAGTCTCTGGACACCTGCATGTCGCCAGCCCCAGGCGGCTGGTTGTCAAGCT	240
QY	241	GGGGCAGGCCCCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT	300
Db	241	GGGGCAGGCCCCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT	300
QY	301	GAGGTGAATGGCCGCGCAGGTACCTGGATGGAGAGACCCTTAAACCCTTGAAGGCTCTG	360
Db	301	GAGGTGAATGGCCGCGCAGGTACCTGGATGGAGAGACCCTTAAACCCTTGAAGGCTCTG	360
QY	361	TGCCGCTGTGATGACGGGTGGCTTACCTGCTGCCGCTGTGCAGTGAAGATGTCGGCTG	420
Db	361	TGCCGCTGTGATGACGGGTGGCTTACCTGCTGCCGCTGTGCAGTGAAGATGTCGGCTG	420
QY	421	CCCAGCTGGGACTGCCCCACGCCCCCAAGAGAATACAGGTGCGCAGGAAGTCTGCCCCGAG	480
Db	421	CCCAGCTGGGACTGCCCCACGCCCCCAAGAGAATACAGGTGCGCAGGAAGTCTGCCCCGAG	480
QY	481	TGGGTATGTGACCAAGGAGTGACACCCGCCGATCCAGCGCTCCACCGCGCCCAAGACACCAA	540
Db	481	TGGGTATGTGACCAAGGAGTGACACCCGCCGATCCAGCGCTCCACCGCGCCCAAGACACCAA	540
QY	541	CTTCTGCCCTTGTCACTCCCTGCTCTGCTGATGCTCTTGTCCAAATTTGAGCACAGCC	600
Db	541	CTTCTGCCCTTGTCACTCCCTGCTCTGCTGATGCTCTTGTCCAAATTTGAGCACAGCC	600
QY	601	TGGGGCCCCCTGTCAACACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAAC	660
Db	601	TGGGGCCCCCTGTCAACACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAAC	660
QY	661	CGATTCTGCCCAACTGGAGATCCAAAGCCGCTGTGTCTGCGCAGACCCCTGCTGGCAGCC	720
Db	661	CGATTCTGCCCAACTGGAGATCCAAAGCCGCTGTGTCTGCGCAGACCCCTGCTGGCAGCC	720
QY	721	AGGAGCCACAGCTCATGGAACAGTCTTTCTAA 753	
Db	721	AGGAGCCACAGCTCATGGAACAGTCTTTCTAA 753	

RESULT 2
AAZ07516

AC AAZ07516;

DT 26-NOV-1999 (first entry)

Rat HICP polypeptide encoding cDNA.

KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss; cell-associated activity; fibroblast activity

XX
CS
Rating

05 Rattus sp.

PN W09947556-A2.

PD 23-SEP-1999.

PF 18-MAR-1999; 99WO-US05999

PR 19-MAR-1998; 98US-0044273.

PA (TUFT) TUFTS COLLEGE.

PI Castellet JJ;

DR WPI; 1999-562060/47.

DR P-PSDB; AAY27434.

PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein used in methods to identify modulators or in diagnostic applications

PS Claim 2; Fig 1; 108pp; English.

This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein. Agents that stimulate or inhibit HICP protein activity or expression, antisense HICP nucleic acid molecules and HICP antibodies, can be used to modulate cell-associated activity. HICP modulators can be used to treat disorders characterized by aberrant HICP protein activity or expression. Probes capable of hybridizing to HICP mRNA or antibodies specific for HICP can be used to detect HICP activity in a biological sample. HICP can be used to treat disorders, such as a cardiovascular or fibrotic disorder, characterized by aberrant cell proliferation.

Sequence 1708. BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Query Match	100.0%;	Score 753;	DB 20;	Length 1708;
Best Local Similarity	100.0%;	Pred. No. 3.2e-193;		
Matches 753; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	I	ATGAGGGGCAGCCCCACTGATCCATCTTCGTGGCCACCTTCTCCTCTCGCCTCTCTCAATG	60
Db	249	ATGAGGGGCAGCCCCACTGATCCATCTTCGTGGCCACCTTCTCCTCTCGCCTCTCTCAATG	308
QY	61	GTTGTGCCCCAGCTGTGCCGGACACCCTGTACCCTGTGGACACACCACCAGTGCCCA	120
Db	309	GTTGTGCCCCAGCTGTGCCGGACACCCTGTACCCTGTGGACACACCACCAGTGCCCA	368
QY	121	CAGGGGGTACCCCCTGGTGTGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGGCTG	180
Db	369	CAGGGGGTACCCCCTGGTGTGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGGCTG	428
QY	181	GGGGAGTCTCTGCGACCACCTGCATGTCGTGCCAGCCCCAGCCAGGGGCTGTGTTGTCAACCT	240
Db	429	GGGGAGTCTCTGCGACCACCTGCATGTCGTGCCAGCCCCAGCCAGGGGCTGTGTTGTCAACCT	488
QY	241	GGGGCAGGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGAATGAGGATGACGTAAGCTGT	300
Db	489	GGGGCAGGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGAATGAGGATGACGTAAGCTGT	548
QY	301	GAGGTGAATGCCCGCAGGTACCTGGATGGAGAGAACCTTTAAACCCAATTGCAGGGCTCCTG	360

Db 549 GAGGTGAATGGCCGACGGTACCTGGATGGAGAGACCTTTAAACCCAATTGCAGGGTCTTG 608
QY 361 TGCCGCTGTATGACGGTGGCTTTCACCTGCTGCCGCTGTGCAGTGAGGATGTGGGCTG 420
Db 609 TGCCGCTGTATGACGGTGGCTTTCACCTGCTGCCGCTGTGCAGTGAGGATGTGGGCTG 668
QY 421 CCCAGCTGGAGTGGCCACGCCCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCCGAG 480
Db 669 CCCAGCTGGAGTGGCCACGCCCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCCGAG 728
QY 481 TGGGTATGTGACCAGGAGTGAACACGGCGATCCAGCGCTCCACGGCGCAAGACACCAA 540
Db 729 TGGGTATGTGACCAGGAGTGAACACGGCGATCCAGCGCTCCACGGCGCAAGACACCAA 788
QY 541 CTTTCTGCCCCTGTCACTCCTGCTCTGCTGATGCTCCTTGTCCAAATTGGAGCACAGCC 600
Db 789 CTTTCTGCCCCTGTCACTCCTGCTCTGCTGATGCTCCTTGTCCAAATTGGAGCACAGCC 848
QY 601 TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAC 660
Db 849 TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAC 908
QY 661 CGATTCTGCCAAGTGGAGATCCAAACGCCGCTGTGTCTGCCAGAACCCCTGCTGGCAGCC 720
Db 909 CGATTCTGCCAAGTGGAGATCCAAACGCCGCTGTGTCTGCCAGAACCCCTGCTGGCAGCC 968
QY 721 AGGAGCCACAGCTCATGGAACAGTCTTTCTAA 753
Db 969 AGGAGCCACAGCTCATGGAACAGTCTTTCTAA 1001

RESULT 3
AAZ07521
ID AAZ07521 standard; cDNA; 681 BP.

XX AAZ07521;

DT 26-NOV-1999 (first entry)

DE Rat HICP mature polypeptide coding sequence.

KW Heparin-induced CN-like protein; HICP; cell-associated activity; ss; cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.

OS Rattus sp.

PN W09947556-A2.

PD 23-SEP-1999.

PF 18-MAR-1999; 99WO-US05999.

PR 19-MAR-1998; 98US-0044273.

PA (TUFT) TUFTS COLLEGE.

PI Castellot JJ;

DR WPI; 1999-562060/47.

DR P-PSDB; AAY27440.

PT Nucleic acid sequences encoding rat heparin-induced CN-like protein,
PT used in methods to identify modulators or in diagnostic applications
XX
PS Disclosure; Fig 2; 108bp; English.

CC The invention provides a rat heparin-induced CN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological

CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents the coding sequence of rat HICP mature
CC polypeptide.

XX
SQ Sequence 681 BP; 121 A; 211 C; 215 G; 134 T; 0 other;

Query Match 90.4%; Score 681; DB 20; Length 681;
Best Local Similarity 100.0%; Pred. No. 6.4e-174;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 CAGCTGTGCCGACACCCCTGTACCTGTCTTGAGACACACCCAGTGCCACAGGGGTA 129
Db 1 CAGCTGTGCCGACACCCCTGTACCTGTCTTGAGACACACCCAGTGCCACAGGGGTA 60
QY 130 CCCCTGTGCTGATGGCTGTGGCTGCTGTAAGTGTGTGCACGAGGCTGGGGAGTCC 189
Db 61 CCCCTGTGCTGATGGCTGTGGCTGCTGTAAGTGTGTGCACGAGGCTGGGGAGTCC 120
QY 190 TGGCACCACCTGCATGTCTGCAACCCACAGCCAGGGCCCTGTTGTCAAGCTGGGCAAGC 249
Db 121 TGGCACCACCTGCATGTCTGCAACCCACAGCCAGGGCCCTGTTGTCAAGCTGGGCAAGC 180
QY 250 CCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT 309
Db 181 CCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAAT 240
QY 310 GGCCGACAGGTACCTGGATGAGAGACCTTTAAACCCAAATTGCAGGGTCTGTGCCGCTGT 369
Db 241 GGCCGACAGGTACCTGGATGAGAGACCTTTAAACCCAAATTGCAGGGTCTGTGCCGCTGT 300
QY 370 GATGACGGTGGCTTACCTGCTGCCCTGCCGCTGTGCAGTGAGGATGTGCGGCTGCCAGCTGG 429
Db 301 GATGACGGTGGCTTACCTGCTGCCCTGCCGCTGTGCAGTGAGGATGTGCGGCTGCCAGCTGG 360
QY 430 GACTGCCCCACGCCCCAAGAGATAAGTGCCACAGGAAAGTGTGCCCCGAGTGGTATGT 489
Db 361 GACTGCCCCACGCCCCAAGAGATAAGTGCCACAGGAAAGTGTGCCCCGAGTGGTATGT 420
QY 490 GACCAGGAGTGAACACCGCGCATCCAGCGCTCCACGCGCGCAAGACACCAACTTTCGCC 549
Db 421 GACCAGGAGTGAACACCGCGCATCCAGCGCTCCACGCGCGCAAGACACCAACTTTCGCC 480
QY 550 CTTGTCACTCCTGCCCTGTGTGATGCTCCTTGTCCAAATTGGAGCACAGCCTGGGGCCCC 609
Db 481 CTTGTCACTCCTGCCCTGTGTGATGCTCCTTGTCCAAATTGGAGCACAGCCTGGGGCCCC 540
QY 610 TGTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAACAGAACCGATTCTGC 669
Db 541 TGTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAACAGAACCGATTCTGC 600
QY 670 CAACGTGAGATCCAAACGCGCTGTGTCTGCCAGACCCCTGCTGGCAGCCAGAGCCAC 729
Db 601 CAACGTGAGATCCAAACGCGCTGTGTCTGCCAGACCCCTGCTGGCAGCCAGAGCCAC 660
QY 730 AGCTCATGGAACAGTGTTC 750
Db 661 AGCTCATGGAACAGTGTTC 681

RESULT 4
AAX76488
ID AAX76488 standard; DNA; 1734 BP.

XX AAX76488;

DT 06-AUG-1999 (first entry)

DE Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;

KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX Mus sp.
XX OS
XX PN W09921998-A1.
XX PD 06-MAY-1999.
XX PE 29-OCT-1998; 98WO-US22991.
XX PR 14-APR-1998; 98US-0081695.
XX PR 29-OCT-1997; 97US-0063704.
XX PR 03-FEB-1998; 98US-0073612.
XX PA (GETH) GENENTECH INC.
XX PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX DR WPI; 1999-337420/28.
XX DR P-PSDB; AAY17651.
XX PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX PS Example 2; Page 178-179; 284pp; English.
XX CC The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoeleic disorders, haematopolesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX SQ Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;
QY Query Match 87.5%; Score 659; DB 20; length 1734;
Best Local Similarity 93.0%; Pred. No. 7.4e-168;
Matches 702; Conservative 0; Mismatches 50; Indels 3; Gaps 1;
QY 1 ATGAGGGGACGCCACTGATCATCTTCTGGCCACTTCCCTCTGCTCTCAATG 60
DB 257 ATGAGGGGCAACCCACTGATCATCTTCTGGCCATTCCCTCTGCAATCTCTCATG 316
QY 61 GTGTGTGCCCAAGCTGTGCCGACACCCCTGTACTGTCTTGGACACACCCCAAGTCCCA 120
DB 317 GTGTATTCCCAAGCTGTGCCGACACCCCTGTGCTGTCTTGGACACACCCCAAGTCCCA 376
QY 121 CAGGGGGTACCCCTGTGTGTGATGGCTGTGGCTGTGTAAGTGTGTGACGAGGCTG 180
DB 377 CCGGGGGGTACCCCTGTGTGTGATGGCTGTGGCTGTGTAAGTGTGTGACGAGGCTG 436
QY 181 GGGAGTCTCTGACACACCTGATGTCTGACACCCACAGGGCCTGTGTTGTACGCT 240
DB 437 GGGAGTCTCTGACACACCTGATGTCTGACACCCACAGGGCCTGTGTTGTACGCT 496
QY 241 GGGGACAGGCCCTGTGGCGGCATGGGGCTGTGTCTCTTGGATGAGGATGACGCTAGCTGT 300

DB 497 GGGGACAGGCCCCAGTGGCCGTGGTGTCTGTGCTCTTCCGAAGAGATGACGGGAGCTGT 556
QY 301 GAGGTGAATGCGCCGACAGTACTGTGATGGAGAGACCTTTAAACCAATTGCAGGCTCTG 360
DB 557 GAGGTGAATGCGCCGACAGTACTGTGATGGGAGAGACCTTTAAACCAATTGCAGGCTTTTG 616
QY 361 TGCCGCTGTGATGACGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 617 TGCCGCTGTGATGACGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
QY 421 CCCAGCTGGGACTGCCACAGCCCCAAGAGATACAGTGGCCAGAAAGTGTGCCCGAG 480
DB 677 CCCAGCTGGGACTGCCACAGCCCCAAGAGATACAGTGGCCAGAAAGTGTGCCCGAG 736
QY 481 TGGGTATGTGACCAAGGAGTGA--CACCAGGATCCAGCCGCTCCACGGCCGACAGACAC 537
DB 737 TGGGTGTGTGACCAAGGAGTGA--CACCAGGATCCAGCCGCTCCACGGCCGACAGACAC 796
QY 538 CACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597
DB 797 CACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
QY 598 GCGTGGGGCCCTGCTCAACACCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 657
DB 857 GCGTGGGGCCCTGCTCAACACCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 916
QY 658 AACCGATTCTGCCAAGTGAATCCAAACGCCGCTGTGTGCTGCCAGACCTGCTGGCA 717
DB 917 AACCGATTCTGCCAAGTGAATCCAAACGCCGCTGTGTGCTGCCAGACCTGCTGGCA 976
QY 718 GCCAGGAGCCACAGCTCATGGAACAGTGTCTTA 752
DB 977 TCCAGGAGCCACAGGCTCATGGAACAGTGTCTTA 1011
RESULT 5
AAX76489/C
ID AAX76489 standard; DNA; 753 BP.
XX AAX76489;
XX AC
XX DT 06-AUG-1999 (first entry)
XX DE Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18.
XX KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopolesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX OS Mus sp.
XX PN W09921998-A1.
XX PD 06-MAY-1999.
XX PE 29-OCT-1998; 98WO-US22991.
XX PR 14-APR-1998; 98US-0081695.
XX PR 29-OCT-1997; 97US-0063704.
XX PR 03-FEB-1998; 98US-0073612.
XX PA (GETH) GENENTECH INC.
XX PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
XX PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX DR WPI; 1999-337420/28.

PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
XX Disclosure; Page 179-180; 284pp; English.
XX
XX The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
XX Sequence 753 BP; 159 A; 229 C; 238 G; 127 T; 0 other;
SQ

Query Match	87.3%;	Score 657;	DB 20;	Length 753;
Best Local Similarity	93.0%;	Pred. No. 2e-167;		
Matches 700; Conservative	0;	Mismatches 50;	Indels 3;	Gaps 1;

QY	1	ATGAGGGGCGACGCCACTGATTCATCTTCTGGCCACTTCCTTCCTCTGCTCTCTCAATG	60
Db	753	ATGAGGGGCGAACCCACTGATTCATCTTCTGGCCATTTCTTCCTTCCTCTGCAATG	694
QY	61	GTGTGTGCCAGCTGTGCCGACACCCTGTACCTGTCTTGGACACCACCCCAGTGCCCA	120
Db	693	GTGTATTCAGAGCTGTGCCAGCACCCCTGTGCCCTGTCTTGGACACCACCCCAGTGCCCA	634
QY	121	CAGGGGGTACCCCTGTGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGCACGGAGGCTG	180
Db	633	CCGGGGGTACCCCTGTGTGTGGATGGCTGTGGCTGTGTCAAGTGTGTGCACGGAGGCTG	574
QY	181	GGGGAGTCTTGCAGCACCTGCATGTCTGCCAGCCCCAGCCAGGGCCGTGTGTCAAGCT	240
Db	573	GGGGAGTCTTGCAGCACCTGCATGTCTGCCAGCCCCAGCCAGGGCCGTGTGTCAAGCT	514
QY	241	GGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT	300
Db	513	GGGGCAGGCCCTGAGTGGCCGTGTGTCTGTGTGCTCTTGAAGAGATGACGGAGCTGT	454
QY	301	GAGGTGAATGGCCGCGAGTACCTGGATGGAGAGACCTTTAAACCAATTGCAGGGTCTG	360
Db	453	GAGGTGAATGGCCGCGAGTACCTGGATGGGGAGACCTTTAAACCAATTGCAGGGTTTG	394
QY	361	TGCCGCTGTATGACGGTGTCTTCACTTGCCCTGCGCTGTGCAGTGAAGATGTGCGGCTG	420
Db	393	TGCCGCTGTATGACGGTGTCTTCACTTGCCCTGCGCTGTGCAGTGAAGATGTGCGGCTG	334
QY	421	CCCAGCTGGAGCTGCCACGCCCCAAGAGAATACAGGTGCCAGAGAAAGTCTGCCCCGAG	480
Db	333	CCCAGCTGGAGCTGCCACGCCCCAAGAGAATACAGGTGCCAGAGAAAGTCTGCCCCGAG	274
QY	481	TGGGTATGTGACCAGGGAGTGA--CACCGGCGATCCAGCGCTCCAGCGCCAGAGACAC	537
Db	273	TGGGTGTGTACCAAGGACGTGATGCAGCCGGCAATCCAGCCCTCTCAAGCCCCAAGGACAC	214
QY	538	CAACTTCTGCCCCCTGTCTACTCTGCTGTCTGTATGCTCCTTGTCCAAATTGGAGCACA	597
Db	213	CAACTTCTGCCCCCTGTCTACTCTGCAATCTGCAATGTGCCGATGGCCCCCTGTCCAAACTGGAGCACA	154
QY	598	GCCTGGGGCCCTGTCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACAG	657

Db	153	GCCTGGGGCCCTGCTCAACCACCCTGTGGGTTGGGCATAGCCACCAGTATCCAACCAG	94
OY	658	AACCGATTCTGCCAACTGGAGATCCAAAGCCGCGCTGTGTCTGCCACAGCCCTGCCTGGCA	717
Db	93	AACCGATTCTGCCAACTGGAGATCCAGCGCTGCGCTGTGTCTGTCCAGACCCTGCCTGGCA	34
OY	718	GCCAGGAGCCACAGCTCATGGAACAGTGCCTTC	750
Db	33	TCCAGGAGCCACGGCTCATGGAACAGTGCCTTC	1

RESULT 6
AAX28435

XX AAX28435;
AC
XX 22-JUN-1999 (first entry)
DT
XX
DE EGF-like homologue PRO261 coding sequence.

Query Match	67.8%;	Score 510.4;	DB 20;	Length 1257;
Best Local Similarity	79.9%;	pred. No. 8.2e-128;		
Matches 601; Conservative	0;	Mismatches 151;	Indels 0;	Gaps 0;

QY	481	TGGGTATGTGACCGAGGAGTGTACACCGGGCGATCCAGCGCTCCACCGGGCGCAGGACCA	540
Db	490	TGGGTGTGCGGCCAAGGAGGGGAGATGGGGACCCAGCCCTTCACAGCCCAAGGACCCAG	549
QY	541	CTTCTGCCCTTGTACATCTGCGCTGTGCTGATGCTCCTTGTCCAAATTGAGCACAGCC	600
Db	550	TTTCTGGCCTTGTCTCTTCCCTGCCCTTGCTGTCCCTGCGCCAGAAATGGAGCACGGCC	609
QY	601	TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAAC	660
Db	610	TGGGACCCCTGCTCGACCACTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAGAAC	669
QY	661	CGATTCTGCCAACTGGAGATCCAACGCCGCCCTGTGTCTGCCCCAGACCCCTGCCCTGGCAGCC	720
Db	670	CGCTTCTGCCGACTGGAGACCCAGCGCCGCTGTGCCCTGTCCAAGGCCCTGCCCCACCCCTCC	729
QY	721	AGGAGCCACAGCTCATGGAACAGTGTCTTCTA	752
Db	730	AGGGGTCCGAGTCCACAAACAGTGTCTTCTA	761

RESULT 8
AAS21403
ID AAS21403 standard; cDNA; 1266 BP.

AC AAS21403;

DT 24-OCT-2001 (first entry)

DE Human cDNA sequence encoding for PRO261 polypeptide.

KM Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KM breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 KM cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KM adipocyte; A-peptide; factor VIIA; gene therapy; ss.

OS Homo sapiens.

PN WO200140466-A2.

PD 07-JUN-2001.

PF 01-DEC-2000; 2000WO-US32678.

PR 01-DEC-1999; 99WO-US28301.

PR 02-DEC-1999; 99WO-US28551.

PR 02-DEC-1999; 99WO-US28565.

PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US309999.

PR 06-JAN-2000; 2000WO-US00277.

PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04342.

PR 24-FEB-2000; 2000WO-US04914.

PR 01-MAR-2000; 2000WO-US05601.

PR 21-MAR-2000; 2000WO-US07532.

PR 17-MAY-2000; 2000WO-US13705.

PR 30-MAY-2000; 2000WO-US14941.

PR 10-NOV-2000; 2000WO-US30873.

PA (GETH) GENENTECH INC.

PA (GETH) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W,
PI Geritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI: 2001-408281/43.
DR P-PSDB; AAU12331.

PT	Isolated, secretory and transmembrane PRO polypeptide used to detect
PT	other PRO polypeptides, link bioactive molecules to cells expressing
PT	PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT	lung, breast, prostate, cervical -

PS Claim 3; Fig 319; 813pp; English.

CC AAS21244-AAS21518 encode for novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIA. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.

Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match	67.8%;	Score 510.4;	DB 22;	Length 1266;
Best Local Similarity	79.9%;	Pred. No. 8.3e-128;		
Matches 601; Conservative	0;	Mismatches 151;	Indels 0;	Gaps 0

QY	1	ATGAGGGGCGACCCACTGATCCATCTTCTGGCCACTTCCTCTCTGCTCTCTCAATG	60
Db	10	ATGAGAGGCACACCGAAGACCACACTCTCTGGCCCTTCCTCTCTGCTCTCAAG	69
QY	61	GTCGTGTCCAGCTGTGCGCGGACACCCCTGATCCTGTCTTGGACACACCCAGTGCCCA	120
Db	70	GTGCGTACCAGCTGTGCCCCGACACCACTGACCTGCCCCCTGGCACCTCCCGATGCCCG	129
QY	121	CAGGGGCTACCCCTGTGTCTGATGGCTGTGGCTGCTGTAAAGTGTGCACGAGGCTG	180
Db	130	CTGGGAGTACCCCTGTGTCTGATGGCTGTGGCTGCTGCGGGTATGTGCACGCGGGCTG	189
QY	181	GGGGAGTCTTGCGACCACTGCATGCTGCGACCCCGACGCCAGGGCCTGTTTGCAGCCT	240
Db	190	GGGGAGCCCTGCGACCACTCCACGTCGCGACGCCAGCCAGGGCCTGTGTCGCCAGCCC	249
QY	241	GGGGCAGGCCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT	300
Db	250	GGGGCAGGACCCGGTGGCCGGGGGGCCCTGTGCCCTTGGCAGAGGACGACAGCAGCTGT	309
QY	301	GAGGTGAATGGCCCGCAGGTACCTGATGGAGAGACCTTTAAACCAATTGCAGGGTCTCG	360
Db	310	GAGGTGAACGGCCCGCCTGTATCGGGAAGGGAGAGACCTTCCAGCCCCCACTGCAGCATCCGC	369
QY	361	TGCCGCTGTGATGACGGGTGGCTTCACCTGCGCTGCCGCTGTGCAGTGAGGATGTGGGCTG	420
Db	370	TGCCGCTGCCAGGACGGCGGCTTCACCTGCTGCCGCTGTGCAGCGAGGATGTGGGCTG	429
QY	421	CCGAGCTGGGACTGCCACAGCCCCCAAGAGAATACAGGTGCCAGGAAGTGTGCCCCGAG	480
Db	430	CCGAGCTGGGACTGCCCCCCACCCAGAGGGGTGAGGTCCTGGGCAAGTGTGCCCCTGAG	489

```
QY 481 TGGTATGTGACCGAGGAGTGACACCGCGATCCAGGGCTCCACGGCGCAAGACACCAA 540
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 TGGGTGTGCGGCCAAGAGAGGGGAGTGGGGACCCAGGCCCTTCCAGCCCAAGACCCAG 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 CTTTTCGCCCTGTGCACTCTGCTCTGCTGATGCTCTGTGCAATTTGAGACACAGCC 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 TTTTCTGCGCTTGTCTCTTCCCTGCCCTGTGCTGCTGCGCCGCAATGAGACAGGCC 609
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGGCACCCGAGTGTCCAACAGAAC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 TGGGGACCTGCTCGACCACTGTGGGCTGGGCATAGGCACCCGCGTGTCCAACAGAAC 669
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 CGATTTCGCCCACTGGAGATCCACACCGCCGCTGTGCTGCCCCAGACCCCTGCCCTGGCAGCC 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 670 CGCTTTCGCCGACTGGAGACCCAGCGCCGCTGTGCTGCTGCTCAGGCGCCCTGCCACCCCTCC 729
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 AGGAGCCACAGCTCATGGAACAGTCTTCTA 752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 730 AGGGGTGCGAGTCCACAAAACAGTGCCTTCTA 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 9

AAF60368 standard; cDNA; 1266 BP.

AAF60368;

27-APR-2001 (first entry)

PRO261 coding sequence.

Cytostatic; PRO protein; tumour; cancer; ss.

Homo sapiens.

WO200105836-A1.

25-JAN-2001.

20-DEC-1999; 99WO-US30999.

20-JUL-1999; 99US-0144758.

26-JUL-1999; 99US-0145698.

08-SEP-1999; 99WO-US20594.

13-SEP-1999; 99WO-US20944.

15-SEP-1999; 99WO-US21090.

05-OCT-1999; 99WO-US23089.

29-NOV-1999; 99WO-US28214.

30-NOV-1999; 99WO-US28313.

02-DEC-1999; 99WO-US28564.

(GETH) GENENTECH INC.

Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;

WPI; 2001-091968/10.

P-PSDB; AAB68598.

New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,

useful for diagnosing and treating cancers -

Claim 50; Fig 13; 196pp; English.

The present invention relates to PRO proteins and coding sequences. The present sequence is the coding sequence for one such PRO protein. It was found that the PRO genes are amplified in the genome of tumour cells. The gene amplification is expected to be associated with the overexpression of the gene product and contributes to tumorigenesis. Therefore, antagonists of PRO proteins are useful for the treatment of benign or malignant tumours, leukemias, lymphoid malignancies and other disorders such as neuronal, glial, astrocytic, hypothalamic, glandular, epithelial, inflammatory and immunologic disorders.

XX SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match

Best Local Similarity 67.8%; Score 510.4; DB 22; Length 1266; Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```
QY 1 ATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCCTCTGCTCTCTCAATG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10 ATGAGAGGACACACCGAAGACCCACCTCTGCGCTTCTCCCTCTGCTCTCAAG 69
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GTGTGCCCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTTGACACCAACCCAGTCCCA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70 GTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGCGCACCTCCCGATGCCCG 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 CAGGGGTACCCCTGTGCTGTGATGCTGTGCTGCTGCTGTAAGTGTGTGACGAGGCTG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 CTGGAGTACCCCTGTGCTGTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GGGAGTCTCGACCACTGACCTGATGCTGCGACCCAGCCAGCGGCTGTTGTAGCCT 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 GGGAGCCCTGCGACCACTGACCTGATGCTGCGACCCAGCCAGCGGCTGTTGTAGCCT 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 GGGGAGCCCTGCGGCGCATGGGCTGTGTCTCTTGGATGAGATGACGCTAGCTGT 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 GGGGAGGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 GAGGTGAATGCCCCGACAGTACCTGATGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 310 GAGGTGAACGCGCGCTGTATCGGGAAGGGAGAGACCTTCCAGCCCACTGACATCCGC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 TGCCGCTGATGACGAGGTGCGCTTACCTGCTGCCGCTGTGCAATGAGGATGTGCGGCTG 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 TGCCGCTGACGAGGACGCGGCTTACCTGCTGCCGCTGTGCAATGAGGATGTGCGGCTG 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 CCCAGCTGGAGCTGCCACGCGCCCAAGAGATACAGGTGCCAGGAAAGTCTGCCCGAG 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 CCCAGCTGGAGCTGCCACGCGCCCAAGAGGAGGTGAGGCTCTGGGCAAGTCTGCCCGAG 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 TGGTATGTGACGAGGAGTGACACCGCGATCCAGCGCTCCACGGCGCAAGACACCAA 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 490 TGGGTGTGCGGCCAAGAGAGGGGAGCTGGGACCCAGCCCTTCCAGCCCAAGACCCAG 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 CTTTTCGCCCTGTGCACTCTGCTCTGCTGATGCTCTGTGCAATTTGAGACACAGCC 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 TTTTCTGCGCTTGTCTCTTCCCTGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGGCACCCGAGTGTCCAACAGAAC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 TGGGGACCTGCTCGACCACTGTGGGCTGGGCATAGGCACCCGCGTGTCCAACAGAAC 669
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 CGATTTCGCCCACTGGAGATCCACACCGCCCTGTGCTGCCCCAGACCCCTGCCCTGGCAGCC 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 670 CGCTTTCGCCGACTGGAGACCCAGCGCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 AGGAGCCACAGCTCATGGAACAGTGTCTTCTA 752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 730 AGGGGTGCGAGTCCACAAAACAGTGCCTTCTA 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 10

AAC97451 standard; cDNA; 1266 BP.

AAC97451;

28-FEB-2001 (first entry)

Human angiogenesis-associated protein PRO261 cDNA, SEQ ID NO:107.

Human; angiogenesis-associated protein; PRO; endothelial cell growth;

cardiac hypertrophy; cardiovascular disorder; endothelial disorder;

angiogenic disorder; atherosclerosis; osteoporosis; hypertension;

KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
KW Crohn's disease; psoriasis; endometriosiis; ulcer; wound healing; cancer;
KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
KW gene therapy; transgenic animal; ss.

OS Homo sapiens.

PN WO200053753-A2.

XX 14-SEP-2000.

XX 05-JAN-2000; 2000WO-US00219.

XX 08-MAR-1999; 99WO-US05028.

PR 12-MAR-1999; 99US-0123957.

PR 14-MAY-1999; 99US-0134287.

PR 02-JUN-1999; 99WO-US12252.

PR 23-JUN-1999; 99US-0141037.

PR 20-JUL-1999; 99US-0144758.

PR 26-JUL-1999; 99US-0145698.

PR 01-SEP-1999; 99WO-US20111.

PR 08-SEP-1999; 99WO-US20594.

PR 15-SEP-1999; 99WO-US21090.

PR 15-SEP-1999; 99WO-US21547.

PR 05-OCT-1999; 99WO-US23089.

PR 30-NOV-1999; 99WO-US28313.

PR 30-NOV-1999; 99WO-US28409.

PR 02-DEC-1999; 99WO-US28564.

PR 02-DEC-1999; 99WO-US28565.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;

PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;

PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;

XX WPI; 2001-090793/10.

DR P-PSDB; AAB53084.

PT New isolated nucleic acid for producing a PRO polypeptide, analyzing
PT genetic disorders and treating cardiovascular, endothelial or
PT angiogenic disorders, such as atherosclerosis, wounds or cancer -

XX Claim 58; Fig 41; 293pp; English.

XX The invention relates to novel human angiogenesis-associated proteins
CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding
CC PRO proteins. The invention also relates to vectors and host cells.
CC comprising a PRO nucleic acid, the recombinant production of a PRO
CC protein, PRO antibodies specific for a PRO protein, fusion proteins
CC comprising a PRO protein, agonists or antagonists of a PRO protein, and
CC compounds which inhibit the expression of a PRO gene. The invention
CC additionally encompasses methods of identifying modulators of PRO
CC expression or activity; diagnosing a cardiovascular, endothelial or
CC angiogenic disorder, or a susceptibility to such a disorder by detecting
CC mutations in a PRO gene, or the expression level of a PRO gene within a
CC particular tissue; treating a cardiovascular, endothelial or angiogenic
CC disorder via the administration of a PRO protein, PRO nucleic acid, or
CC PRO agonist or antagonist; a retroviral gene therapy vector comprising a
CC PRO nucleic acid; and methods of inhibiting or stimulating endothelial
CC cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the
CC administration of a PRO protein, or an agonist or antagonist thereof.
CC PRO nucleic acids, PRO proteins, antibodies against PRO proteins, PRO
CC agonists and PRO antagonists may be used as therapeutic agents to treat
CC cardiovascular, endothelial or angiogenic disorders, such as
CC atherosclerosis, osteoporosis, myocardial infarction, hypertension,
CC diabetic retinopathy, rheumatoid arthritis, Crohn's disease, psoriasis,
CC endometriosiis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's
CC disease, or stroke. PRO nucleic acids are additionally useful in the
CC recombinant production of PRO proteins, as hybridisation probes to
CC screen libraries to isolate cDNAs with sequence identity to PRO proteins,
CC to map genes encoding PRO proteins, to analyse genetic disorders, and in
CC gene therapy. PRO nucleic acids can also be used to produce transgenic

CC animals useful for the development and screening of potential
CC therapeutic agents. The present sequence represents a cDNA encoding a PRO
CC protein of the invention.

XX SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 67.8%; Score 510.4; DB 22; Length 1266;

Best Local Similarity 79.9%; Pred. No. 8.3e-128;

Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGACGCCCACTGATCCATCTTGCGCACCTTCCTTCTGCTTCTCAATG 60

Db 10 ATGAGAGGACACCGAAGACCCACTCTGCGCTTCTCCCTCTGCTCTCAAG 69

QY 61 GTGTGTGCCCAAGCTGTGCCGGACACCTGTACTGTCTGCTTGACACACCCAGTGCCCA 120

Db 70 GTGCGTACCCAGCTGTGCCCGACACCATGTACTGTGCCCTTGCCACCTCCCGATGCCG 129

QY 121 CAGGGGTACCCCTGTGCTGATGCTGTGCTGCTGTGCTGTAAAGTGTGACGGAGCTG 180

Db 130 CTGGGATACCCCTGTGCTGATGCTGTGCTGCTGCTGCTGCCGGGTATGTGACGGCGCTG 189

QY 181 GGGAGTCTTGCAGCACCTGCATCTGTGCAGACCCAGCCAGGCGCTGTTGTACGCT 240

Db 190 GGGAGCGCTGCGACCACTCCACGCTGTGCGAGCCAGCCAGGCGCTGTTGTACGCT 249

QY 241 GGGGAGCGCTGCGCGCATGGCGCTGTGCTCTGTGATGAGATGACGCTAGCTGT 300

Db 250 GGGGAGGACCGCGGTGCGCGCGCGCGCGCGCTGTGCTCTGTGAGAGGACGACGACGCTGT 309

QY 301 GAGGTATGCGCGGACGATACCTGTGATGAGAGACCTTTAAACCAATTGACGGTCTGT 360

Db 310 GAGGTATGCGCGCGGACGATACCTGTGATGAGAGACCTTTAAACCAATTGACGGTCTGT 369

QY 361 TGCCGCTGTGATGACGGGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

Db 370 TGCCGCTGTGAGAGACGGCGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429

QY 421 CCCAGCTGGGACTGCCACGCGCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCGGAG 480

Db 430 CCCAGCTGGGACTGCCACGCGCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCGGAG 489

QY 481 TGGGTATGTGACGAGGAGTGACCGGCGATCCAGCGCTCCAGCGCGCAAGACACCAA 540

Db 490 TGGGTATGTGACGAGGAGTGACCGGCGATCCAGCGCTCCAGCGCGCAAGACACCAA 549

QY 541 CTTTCTGCCCTTGTCACTCTGCTCTGCTGATGCTCTTGTCCAAATGGAGCACAGCC 600

Db 550 TTTTCTGCCCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 609

QY 601 TGGGCGCGCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAACAGAAC 660

Db 610 TGGGCGCGCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAACAGAAC 669

QY 661 CGATTCTGCCAAGTGAAGATCAAGCGCGCTGTGTCTGCGAGACCCCTGCTGGCAGCC 720

Db 670 CGATTCTGCCAAGTGAAGATCAAGCGCGCTGTGTCTGCGAGACCCCTGCTGGCAGCC 729

QY 721 AGGAGCCACAGCTCATGGAACAGTGTCTTCTA 752

Db 730 AGGAGCCACAGCTCATGGAACAGTGTCTTCTA 761

RESULT 11

AAV29260

ID AAV29260 standard; DNA; 1285 BP.

XX AAV29260;

XX 14-SEP-1998 (first entry)

DE Human connective tissue growth factor-3 gene.

XX

KW Connective tissue growth factor-3; CTGF-3; human; cancer;
KW arthritis; fibrosis; osteoporosis; diagnosis; therapy; ds.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 9..761
FT sig_peptide 9..65
FT mat_peptide 66..758
FT /*tag= a
FT /*tag= b
FT /*tag= c
XX WO9821236-A1.
PN 22-MAY-1998.
XX 08-NOV-1996; 96WO-US17856.
XX 08-NOV-1996; 96WO-US17856.
XX 08-NOV-1996; 96WO-US17856.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Chopra A, Ebner R, Ruben SM;
PI WPI: 1998-297864/26.
DR P-PSDB; AAW37946.
XX
PT Novel human connective tissue growth factor 3 gene - useful for the
PT diagnosis and treatment of e.g. cancer, arthritis, fibrosis,
PT osteoporosis
XX
PS Claim 2; Fig 1A-B; 87pp; English.
XX
XX This nucleotide sequence codes for human connective tissue growth
CC factor-3 (CTGF-3) protein (see AAW37946), a novel member of the
CC growth factor superfamily. It was discovered in a cDNA library
CC derived from human osteoblasts. The gene has also been identified
CC in cDNA libraries from ovary, testis, heart, lung, skeletal muscle,
CC adrenal medulla, adrenal cortex, thymus, prostate, small intestine
CC and colon. A cDNA clone is deposited as ATCC 97756. Also provided
CC are vectors, host cells and recombinant methods for producing
CC CTGF-3 polypeptides. CTGF-3 nucleic acid sequences or their
CC fragments, e.g. primers or probes, can be used to diagnose diseases
CC where CTGF-3 expression is enhanced, e.g. cancer, arthritis,
CC fibrosis or atherosclerosis, or diseases where expression is
CC decreased such as in osteoporosis. Disorders characterised by
CC decreased or increased levels of CTGF-3 can be treated by
CC administering CTGF-3 polypeptides and anti-CTGF-3 antibodies,
CC respectively.
XX
SQ Sequence 1285 BP; 237 A; 418 C; 389 G; 241 T; 0 other;

Query Match 67.8%; Score 510.4; DB 19; Length 1285;
Best Local Similarity 79.9%; Pred. No. 8.3e-128;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGAGCCCACTGATTCATCTCTGGCCACTTCCCTCTGCTCTCAATG 60
DB 9 ATGAGAGGACACCGAAGACCACCTCTGCGCTTCTCCGCTCTGCTCTCAAG 68
QY 61 GTGTGTCCCACTGTGCGGACACCTGTACCTGTCTGTGACACACCCAGTCCCA 120
DB 69 GTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCG 128
QY 121 CAGGGGTACCCCTGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
DB 129 CTGGGATACCCCTGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 188
QY 181 GGGGAGTCTCGGACCACTGCATGTCTGCGAACCAGCCAGGGGCTGTGTCAGCCT 240
DB 189 GGGGAGCCTCGACCACTGCATGTCTGCGAACCAGCCAGGGGCTGTGTCAGCCT 248

QY 241 GGGGAGCCTCGGACCACTGCATGTCTGCTGTCTCTGTGATGAGATGACGCTAGCTGT 300
DB 249 GGGGAGGACCCCGGTGGCCCGGGGGCCCTGTGCTCTGTGGCAGAGACGACAGCAGCTGT 308
QY 301 GAGGTGATGCGCCGACAGGTACCTGATGAGAGACCTTTAAACCAATTCAGGGTCTCTG 360
DB 309 GAGGTGATGCGCCGCGCTGTATGCGAAGGAGAGACCTTCCAGCCCACTGACATCCGC 368
QY 361 TGCCGCTGTGATGACGGGTGCTTCACTGCTCCGCTGTGAGTGAGATGTGCGGCTG 420
DB 369 TGCCGCTGTGAGGACGCGGCTTCACTGCTGCGCTGTGAGTGAGATGTGCGGCTG 428
QY 421 CCCAGCTGGGACTGCCCCACGCCCCAAGAGATACAGGTGCCAGAAAGTGTGCCCGAG 480
DB 429 CCCAGCTGGGACTGCCCCACGCCCCAAGAGATGAGGTGCTGCGCAAGTGTGCCCTGAG 488
QY 481 TGGGTATGTGACCAAGGAGTGACACCGCGCATCCAGCGCTCCAGCGGCAAGACACCA 540
DB 489 TGGGTATGTGACCAAGGAGTGACACCGCGCATCCAGCGCTCCAGCGGCAAGACACCA 548
QY 541 CTTTCTGCGCTGTGCTACCTCCTGCTCTGCTGATGCTCTGTTCCAAATTGAGACACAGCC 600
DB 549 TTTTCTGCGCTGTGCTACCTCCTGCTCTGCTGATGCTCTGTTCCAAATTGAGACACAGCC 608
QY 601 TGGGCGCCCTGCTCAACCACTGTGGGCTGGGATAGCCACCCAGTGTCCAAACAGAAC 660
DB 609 TGGGCGCCCTGCTCAACCACTGTGGGCTGGGATAGCCACCCAGTGTCCAAACAGAAC 668
QY 661 CGATTCTGCCAAGTGGAGATCCAGCCGCGCTGTGCTGTGCCAGACCTGCTGGCAGCC 720
DB 669 CGATTCTGCCAAGTGGAGATCCAGCCGCGCTGTGCTGTGCCAGACCTGCTGGCAGCC 728
QY 721 AGGAGCCACAGCTCATGGAACAGTCTTCTA 752
DB 729 AGGAGCCACAGCTCATGGAACAGTCTTCTA 760

RESULT 12
AAAX76486 standard; DNA; 1293 BP.
XX
XX AAAX76486;
AC
XX 06-AUG-1999 (first entry)
DT
XX
DE Human WISP-2 protein nucleotide sequence SEQ ID NO:13.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Homo sapiens.
XX
PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI: 1999-337420/28.

CC composition of the invention include inhibitors of urokinase-type
CC plasminogen activator (uPA) and matrix metalloproteinase (MMP). The
CC composition is useful for the treatment of chronic damaged tissue, i.e.
CC wounds and dermal ulcers. The present sequence encodes a human
CC CTGF-like protein, and is used to produce the composition of the
CC invention.

SO Sequence 1309 BP; 261 A; 418 C; 387 G; 242 T; 1 other;

Query Match 67.8%; Score 510.4; DB 22; Length 1309;
Best Local Similarity 79.9%; Pred. No. 8.4e-128;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

OY 1 ATGAGGGGCGAGCCCACTGATCATCTTCTGGCCACTTCCCTTCTGCTTCTCAATG 60
DB 7 ATGAGAGGCACACCGAAGACCCTCTGCGCTTCTCTCTCTGCTTCTCAAG 66
OY 61 GTGTGTGCCCAGCTGTGCGGACACCTGTACTGTCTCTGAGACACCAAGCCAGTGCCCA 120
DB 67 GTGCGTACCCAGCTGTGCGCCGACACCATGTACTGCCCCCTGGCCACCTCCCGATGCCCG 126
OY 121 CAGGGGTACCCCTGTGTGCTGATGCTGTGCTGTGCTGTAAAGTGTGTGACAGGAGCTG 180
DB 127 CTGGGAGTACCCCTGTGTGCTGATGCTGTGCTGTGCTGCTGCGGGTATGTGACAGGGGCTG 186
OY 181 GGGAGTCTCTGCGACCACTGCATGTCTGCGACCCACGCGCCAGGGCCTGGTTGTACAGCT 240
DB 187 GGGAGCCCTGCGACCACTGCATGTCTGCGACCCACGCGCCAGGGCCTGGTTGTACAGCT 246
OY 241 GGGGAGGCCCTGCGCGCCATGGGGCTGTGTCTCTTGTGATGAGGATGACGGTAGCTGT 300
DB 247 GGGGAGGACCCCGGTGGCGCGGGGCGCTGTGCTCTTGGCAGAGGACGACAGCAGCAGCTGT 306
OY 301 GAGGTGAATGGCCGAGGTACTGTGATGAGAGAGACCTTTAAACCCATTTGACAGGGTCTG 360
DB 307 GAGGTGAACGGCGCCTGTATGGGGAAGGGAGAGACCTTCCAGCCCACTGACATCCGC 366
OY 361 TGCCGCTGTGATGACGGTGTGCTTCACTGTGCTGCGCTGTGCTGAGTGTGCGGCTG 420
DB 367 TGCCGCTGTGAGAGACGGCGCTTCACTGTGCTGCGCTGTGCTGAGATGTGCGGCTG 426
OY 421 CCCAGCTGGGACTGCCACGCGCCCAAGAGATATACAGGTGCGAGAAAGTGTGCCCGAG 480
DB 427 CCCAGCTGGGACTGCCACGCGCCCAAGAGGTGAGGTCTGCGGCAAGTGTGCCCTGAG 486
OY 481 TGGGTATGTGACAGGAGGTGACACCGCGGATCCAGCGCTCCACGCGCGCAGAGACCA 540
DB 487 TGGGTGTGCGGCGCAAGAGGGGAGCTGGGGAGCCAGCCCTTCCAGCCCAAGAGACCC 546
OY 541 CTTCCTGCCCCCTGTCACTCTCTGCTGTGATGTCTCTTGTCCAAATTGGAGCACAGCC 600
DB 547 TTTTCTGGCCTTGTCTCTTCTTCCCTGCCCCCTGTGCTGCCCAAGATGAGACAGGCC 606
OY 601 TGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGACCACCGAGTGTCCAACAGAAC 660
DB 607 TGGGAGCCTGTCTGACCACTGTGGGCTGGGCATGGCCACCGGGGTGTCCAACAGAAC 666
OY 661 CGATTCTGCCAAGTGAAGTCAACGCGCCCTGTGTCTGCGCAGAACCTGCTGGCAGCC 720
DB 667 CGCTTCTGCGACTGAGAGACCCAGCGCCGCTGTGCTGTCTGCAAGCCCTGCCACCCCT 726
OY 721 AGGAGCCACAGCTCATGGAACAGTCTTCTA 752
DB 727 AGGGGTCCAGTCCACAAACAGTGCCTTCTA 758

RESULT 14
AA16595
ID AAX16595 standard; cDNA; 1522 BP.
XX AAX16595;
XX 29-APR-1999 (first entry)

XX Human growth factor-like protein encoding cDNA.
DE
XX Human; growth factor-like protein; HGFLP; fisp-12; CTGF; PDGF family;
KW connective tissue growth factor; platelet-derived growth factor; keloid;
KW connective tissue disorder; cancer; ankylosing spondylitis; scleroderma;
KW atherosclerosis; Dupuytren's contracture; eosinophilic fasciitis;
KW Feltly syndrome; Goodpasture's disease; Hunter syndrome; Hurler syndrome;
KW Marfan syndrome; nodular fasciitis; osteogenesis imperfecta; restenosis;
KW rheumatoid arthritis; systemic lupus erythematosus; ss.

OS Homo sapiens.

FH Key Location/Qualifiers
FT CDS 256..1008
FT /*tag= a

PN W09858063-A1.

PD 23-DEC-1998.

PE 18-JUN-1998; 98WO-US12787.

PR 19-JUN-1997; 97US-0878990.

PA (INCY-) INCYTE PHARM INC.

PI Corley NC, Hillman JL, Lai P;

DR WPI; 1999-080954/07.

DR P-PSDB; AAW94616.

PT New polynucleotide encoding growth factor-like protein - useful for
PT preventing and treating connective tissue disorders and cancer
PS Claim 4; Fig 1; 63pp; English.

XX The present sequence encodes human growth factor-like protein (HGFLP)
CC which has homology to connective tissue growth factor (CTGF) and fisp-12,
CC which are members of the platelet-derived growth factor (PDGF)
CC superfamily of growth factors. Antisense oligonucleotides and HGFLP
CC antagonists can be used to treat or prevent cancer, and connective
CC tissue disorders including ankylosing spondylitis, atherosclerosis,
CC Dupuytren's contracture, eosinophilic fasciitis, Feltly syndrome,
CC Goodpasture's disease, Hunter syndrome, Hurler syndrome, Marfan
CC syndrome, nodular fasciitis, osteogenesis imperfecta, polyarthritits
CC nodosa, rheumatoid arthritis, scleroderma, systemic lupus erythematosus,
CC and restenosis following angioplasty. HGFLP antibodies and
CC polynucleotides can also be used in diagnostic assays for conditions or
CC diseases characterised by GRFLP expression.

SO Sequence 1522 BP; 279 A; 496 C; 459 G; 288 T; 0 other;

Query Match 67.8%; Score 510.4; DB 20; Length 1522;
Best Local Similarity 79.9%; Pred. No. 8.8e-128;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

OY 1 ATGAGGGGCGAGCCCACTGATCATCTTCTGGCCACTTCCCTTCTGCTTCTCAATG 60
DB 256 ATGAGAGGCACACCGAAGACCACCTCTGCGCTTCTCTCTCTGCTTCTCAAG 315
OY 61 GTGTGTGCCCAGCTGTGCGGACACCTGTACTGTCTTGTGACACCAAGCCAGTGCCCA 120
DB 316 GTGCGTACCCAGCTGTGCGCCGACACCATGTACTGCCCCCTGGCCACCTCCCGATGCCCG 375
OY 121 CAGGGGTACCCCTGTGTGCTGATGCTGTGCTGTGCTGTAAAGTGTGTGACAGGAGCTG 180
DB 376 CTGGGAGTACCCCTGTGTGCTGATGCTGTGCTGTGCTGCGGGTATGTGACAGGGGCTG 435
OY 181 GGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGCGCGCGCTGTGTGTACAGCT 240
DB 436 GGGAGCCCTGCGACCACTGCATGTCTGCGAGCGCCAGCGCGCGCTGTGTGTACAGCCC 495

QY	241	GGGGCAGGGCCCTGGCGGGCCATGGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT	300
Db	496	GGGGCAGGACCCGGTGGCCGGGGGGCCCTGTGCTCTTGGCAGAGAGACGACGACGACTGT	555
QY	301	GAGGTGAATGGCCCGCAGAGTAACCTGGATGGAGAGACCCTTTAAACCAATTGCAGGGTCC	360
Db	556	GAGGTGAACGGCCGCTGTATCGGGGAAGGGAGAGACCTTCAGCCCCACTGCAGCATCCGC	615
QY	361	TGCCGCTGTGATGACGGTGGCTTCACTGCTGCGCTGTGAGTGAAGATGTGCGGCTG	420
Db	616	TGCCGCTGCGAGACGGCGGCTTCACTGCTGCGCTGTGACGCGAGATGTGCGGCTG	675
QY	421	CCGAGCTGGAGCTGCCACGCCCCAAGAGAATACAGGTGCCAGGAAGTGTGCCCCGAG	480
Db	676	CCGAGCTGGAGCTGCCACCCACAGGAGGGTCCAGGTCTGGGCAAGTGTGCCCTGAG	735
QY	481	TGGGTATGTGACCAGGAGTGAACACCGGCGATCCAGCGCTCCACGGGCCAAGGACACCAA	540
Db	736	TGGGTGTGCGGCCAAGGAGGGGAGACTGGGGACCCAGCCCCCTTCAGGCCCAAGGACCCAG	795
QY	541	CTTTCCTGCCCTTGTCACTCCCTGCTGCTGATGCTCTTGTCCAATTGGAGCACAGCC	600
Db	796	TTTTCCTGGCCCTTGTCTCTCCCTGCCCCCTGCTGTCCCTGCCCAGATGGAGCACAGGCC	855
QY	601	TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGGCATAGCCACCCGAGTGTCCAACCAAGAAC	660
Db	856	TGGGGACCCCTGCTCGAACCACTGTGGGCTGGGGCATGGCCACCCGGGTGTCCAACCAAGAAC	915
QY	661	CGATTCTGCCAAGTGGAGATCCAACGCCGCTGTGTCTGCCCCAGAGACCTGTGCGAGCC	720
Db	916	CGCTTCTGCCAGACTGGAGACCCAGCGCCGCTGTGCTGTCCAGAGGCCCTGTGCCACCCCTCC	975
QY	721	AGGAGCCACAGCTCATGGAACAGTGTCTTA 752	
Db	976	AGGGGTGCGAGTCCACAAACAGTGCCTTCTA 1007	

XX	RESULT 15
XX	AAH46952
ID	AAH46952 standard; cDNA; 1337 BP.
XX	
AC	AAH46952;
XX	
DT	25-SEP-2001 (first entry)
XX	
DE	Human secreted protein encoding cDNA (clone Id HBODE48).
XX	
KW	Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW	antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW	neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW	ophthalmological; gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200155430-A1.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01431.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	12-SEP-2000; 2000US-0231968.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI	Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI	Ni J, Ruben SM, Barash SC;
XX	
DR	WPI; 2001-476220/51.
DR	P-PSDB; AAB85542.
XX	

PT 17 isolated nucleic acid molecules encoding human secreted proteins
PT used to preventing, treating or ameliorating a medical condition -
XX
PS Claim 1; page 427-428; 482pp; English.

The invention provides novel human secreted proteins and polynucleotides encoding them. The secreted proteins can be expressed by standard recombinant methodology. The secreted proteins and polynucleotides are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can also be used in diagnosing a pathological condition. The antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. The present sequence represents a human secreted protein encoding cDNA.

Sequence 1337 BP; 257 A; 427 C; 396 G; 252 T; 5 other;

Query Match 67.7%; Score 510; DB 22; Length 1337;

Best Local Similarity	79.8%;	Pred. No. 1.1e-127;	
Matches	600;	Conservative	1; Mismatches 151; Indels 0; Gaps 0

[illegible]

QY 61 GTGTGTGCCCCAGCTGTGCGCGACACCCCTGTACCTGTCTTGACACACCCCAAGTSCCA 120
 ||| || ||||||||| ||||| ||||||| || ||| ||||| |||||
 Db 107 GTGCGTACCCAGCTGTGCCCCGACACCATGTACCTGCCCCCTGGCCACCTCCCCGATGCCCG 166

QY 121 CAGGGGTACCCCTGCTGCTGATGGCTTGGCTGCTGTAAGTGTTGCACGAGGCTG 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 167 CTGGGAGTACCCCTGCTGCTGGATGGCTTGGCTGCTGCCGGGTATGTGCACGGCGGCTG 226

QY 181 GGGAGTCTCGGACCACCTGCATGTCTGCGACCCAGCCAGGGCCTGTTTGTCAGCCT 240
||||| ||||| | | | ||||| . ||||| ||||| | | |||||
Db 227 GGGGAGCCCTGCGACCAACTCCACGTCTGCGAGGCCAGCCAGGGCCTGGTCTGCCAGCCC 286

QY 241 GGGCAGCCCTGGCGGCATGGGGCTGTGTCTCTTGATGAGGATGACGTAACCTGT 300
 ||||| | | | | : | |||| | |||| | |||| | |||| |
 Db 287 GGGCAGGACCCGGTGGMCGGGGGCCCTGTGCTCTTGCGAGAGGACGACAGCACTGT 346

QY 301 GAGTGAATGCGCCGACGGTACCCTGGATGCAGACCTTTAAACCAATTGCAGGGTCCTG 360
||||||| ||||| ||| | ||| ||| ||||| ||||| |||||
Db 347 GAGTGAACGGCGCCGCTGTATCGGAGGGGAGACTTCCAGCCCCACTGCAGCATCCGC 406

QY 361 TGCCTGTGATGACGGTGGCTTCACCTGCCCTGCCCTGTGCAGTGAGGATGTGGCGCTG 420
||||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
dh 407 TGCCCTCCTCCAGAGAGCGCGCTTCACCTGCCCTGCCCTGTGCAGGAGATGTGCCCGCTG 466

421 CCACGCTGGGACTGCCCCACGCCCCCAAGAGATACAGGTGCCAGGAAGTGTGCCCCGAG 480
 467 CCGACGCTGGGACTGCCCCACGCCCCCAAGAGATACAGGTGCCAGGAAGTGTGCCCCGAG 526

[illegible][illegible]

QY	601	TGGGGCCCTGCTCAACCACTGTGGCTGGGCATAGCCA	660
Db	647	TGGGGACCCCTGCTCGACCACTGTGGCTGGGCATAGCCA	706
QY	661	CGATTCTGCCAAGTGGAGATCCAAAGCCGCTGTGTGCG	720
Db	707	CGCTTCTGCCAAGTGGAGATCCAAAGCCGCTGTGTGCG	766
QY	721	AGGAGCCACAGCTCATGGAACAGTGCCTTCTA	752
Db	767	AGGGGTGCGAGTCCACAAACAGTGCCTTCTA	798

Search completed: July 23, 2003, 11:15:20
Job time : 200.398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:55:06 ; Search time 40.9673 Seconds
(without alignments)
5636.882 Million cell updates/sec

Title:	US-10-010-408-3
Perfect score:	753
Sequence:	1 ATGAGGGGACGCCACTGAT.....CATGACACAGTGCCTTCTAA 753

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

```
Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

[illegible]

28	102.8	13.7	1128	2	US-08-459-101A-1	Sequence 1, Appl1
29	98.4	13.1	1062	4	US-09-253-316-3	Sequence 3, Appl1
30	89.2	11.8	1403	4	US-09-182-145-23	Sequence 23, Appl1
31	84	11.2	4214	4	US-09-122-135-1	Sequence 1, Appl1
32	75.4	10.0	1142	4	US-09-253-316-1	Sequence 1, Appl1
33	75.4	10.0	1212	4	US-09-182-145-34	Sequence 34, Appl1
34	75.4	10.0	1212	4	US-09-182-145-35	Sequence 35, Appl1
35	75.4	10.0	1335	4	US-09-182-145-30	Sequence 30, Appl1
36	75.4	10.0	1335	4	US-09-182-145-31	Sequence 31, Appl1
37	65.8	8.7	1101	4	US-09-182-145-29	Sequence 29, Appl1
38	63.4	8.4	693	4	US-09-182-145-24	Sequence 24, Appl1
39	63.4	8.4	1202	4	US-09-182-145-26	Sequence 26, Appl1
40	48.6	6.5	2541	2	US-08-656-393-1	Sequence 1, Appl1
41	41.4	5.5	51	4	US-09-182-145-117	Sequence 117, App
42	39.8	5.3	546	4	US-09-182-145-25	Sequence 28, Appl1
43	39.8	5.3	683	4	US-09-182-145-25	Sequence 25, Appl1
44	39.8	5.3	1183	4	US-09-182-145-27	Sequence 27, Appl1
45	36.6	4.9	4403765	4	US-09-103-840A-2	Sequence 2, Appl1

ALIGNMENTS

```

RESULT 1
US-09-182-145-17
; Sequence 17, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-182-145-17

```

Query Match	87.5%	Score 659;	DB 4;	Length 1734;
Best Local Similarity	93.0%;	Pred. No. 8.5e-161;		
Matches 702; Conservative	0;	Mismatches 50;	Indels 3;	Gaps 1

[illegible]

Db 437 GGGAGTCTCGACACACCTGCAATGTCTGCGACCCAGCCAGGCGCTGTTGTACGCT 496
QY 241 GGGGAGGCCCCCTGGCGCATGGGGCTGTGTCTCTGATGAGGATGACGGTAGCTGT 300
Db 497 GGGGAGGCCCCAGTAGTGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 556
QY 301 GAGTGAATGGCGCAGGTACCTGATGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
Db 557 GAGTGAATGGCGCAGGTACCTGATGAGAGACCTTTAAACCAATTGACAGGCTTTG 616
QY 361 TGCCGCTGTGATGACGGTGGCTTACCTGCTGCGCTGTGTGAGTGAAGATGTGCGGCTG 420
Db 617 TGCCGCTGTGATGACGGTGGCTTACCTGCTGCGCTGTGTGAGTGAAGATGTGCGGCTG 676
QY 421 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCCGAG 480
Db 677 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCCGAG 736
QY 481 TGGGTATGTGACGAGGAGTGA---CACCGCGATCCAGCGCTCCACGCGCAAGACAC 537
Db 737 TGGGTATGTGACGAGGAGTGA---CACCGCGATCCAGCGCTCCACGCGCAAGACAC 796
QY 538 CAACCTTCTGCCCCCTGTGTACCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 597
Db 797 CAACCTTCTGCCCCCTGTGTACCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 856
QY 598 GCGTGGGGCCCCCTGTGTACACCACTGTGGGCTGGGCAAGCCCGAGTGTCCAACCAAG 657
Db 857 GCGTGGGGCCCCCTGTGTACACCACTGTGGGCTGGGCAAGCCCGAGTGTCCAACCAAG 916
QY 658 AACCGATTCTGCCAAGTGGAGATCCAAAGCCCGCTGTGTGTGTGTGTGTGTGTGTGTGT 717
Db 917 AACCGATTCTGCCAAGTGGAGATCCAAAGCCCGCTGTGTGTGTGTGTGTGTGTGTGTGT 976
QY 718 GCCAGAGCCACAGCTCATGGAACAGTCTTCTA 752
Db 977 TCCAGAGCCACAGCTCATGGAACAGTCTTCTA 1011

RESULT 2

US-09-182-145-18/c

; Sequence 18, Application US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P1176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; EARLIER FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 18

; LENGTH: 1734

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-182-145-18

Query Match 87.5%; Score 659; DB 4; Length 1734;
Best Local Similarity 93.0%; Pred. No. 8.5e-161;

Matches 702; Conservative 0; Mismatches 50; Indels 3; Gaps 1;
QY 1 ATGAGGGGACAGCCACTGATCCATCTTCTGTGGCACTTCCCTTCTGTCTCATG 60
Db 1478 ATGAGGGGACAGCCACTGATCCATCTTCTGTGGCACTTCCCTTCTGTCTCATG 1419
QY 61 GTGTGTGCCCCAGCTGTGCCCCGAGACCCCTGTACCTGTCTTGTGAGACACCACCAGTGGCCA 120
Db 1418 GTGTGTGCCCCAGCTGTGCCCCGAGACCCCTGTACCTGTCTTGTGAGACACCACCAGTGGCCA 1359
QY 121 CAGGGGTACCCCTGTGTGCTGTGATGGCTGTGTGCTGTGTGATGAAGTGTGTGACGGAGCTG 180
Db 1358 CCGGGGTACCCCTGTGTGCTGTGATGGCTGTGTGCTGTGTGATGAAGTGTGTGACGGAGCTG 1299
QY 181 GGGAGTCTCGACACACCTGCAATGTCTGCGACCCCAAGGAGGCGCTGTTGTACGCT 240
Db 1298 GGGAGTCTCGACACACCTGCAATGTCTGCGACCCCAAGGAGGCGCTGTTGTACGCT 1239
QY 241 GGGGAGGCCCCCTGGCGCATGGGGCTGTGTGTCTCTTGTGATGAGATGACGGTAGCTGT 300
Db 1238 GGGGAGGCCCCAGTAGTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179
QY 301 GAGTGAATGGCGCAGGTACCTGATGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
Db 1178 GAGTGAATGGCGCAGGTACCTGATGAGAGACCTTTAAACCAATTGACAGGCTTTTG 1119
QY 361 TGCCGCTGTGATGACGGTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 1118 TGCCGCTGTGATGACGGTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1059
QY 421 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCCGAG 480
Db 1058 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCCGAG 999
QY 481 TGGGTATGTGACGAGGAGTGA---CACCGCGATCCAGCGCTCCACGCGCAAGACAC 537
Db 998 TGGGTATGTGACGAGGAGTGA---CACCGCGATCCAGCGCTCCACGCGCAAGACAC 939
QY 538 CAACCTTCTGCCCCCTGTGTACCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 597
Db 938 CAACCTTCTGCCCCCTGTGTACCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 879
QY 598 GCGTGGGGCCCCCTGTGTACCACTGTGGGCTGGGCAAGCCCGAGTGTCCAACCAAG 657
Db 878 GCGTGGGGCCCCCTGTGTACCACTGTGGGCTGGGCAAGCCCGAGTGTCCAACCAAG 819
QY 658 AACCGATTCTGCCAAGTGGAGATCCAAAGCCCGCTGTGTGTGTGTGTGTGTGTGTGTGT 717
Db 818 AACCGATTCTGCCAAGTGGAGATCCAAAGCCCGCTGTGTGTGTGTGTGTGTGTGTGTGT 759
QY 718 GCCAGAGCCACAGCTCATGGAACAGTCTTCTA 752
Db 758 TCCAGAGCCACAGCTCATGGAACAGTCTTCTA 724

RESULT 3

US-09-182-145-13

; Sequence 13, Application US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P1176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 13
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-13

Query Match 67.8%; Score 510.4; DB 4; Length 1293;
Best Local Similarity 79.9%; Pred. No. 1.4e-122;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGAGCCCACTGATCCATCTCTGCGCACTTCCCTCTGCTCTCTCAATG 60
DB 22 ATGAGAGGACACCGAAGACCACCTCTGCGCTTCTCCCTCTGCTCTCTCAAG 81
QY 61 GTGTGCCCCAGCTGTGCGGACACCTGTACTGTCTCTGACACCAAGTGGCCCA 120
DB 82 GTGCGTACCAGCTGTGCCCCAGACCACTGACCTGCCCTGGCCACCTCCCGATGCCG 141
QY 121 CAGGGGTACCCCTGTGTCTGATGGCTGTGGCTGCTGTAAGTGTGCACGAGGCTG 180
DB 142 CTGGAGTACCCCTGTGTCTGATGGCTGTGGCTGCTGCCGGTATGTGCACGGGGCTG 201
QY 181 GGGAGTCCCTGCGACCACTGCATGTCTGCGACCCCAAGCCAGGGCTGTTGTACGCT 240
DB 202 GGGAGGCTTGGACCACTCCACGCTGTGCGACGCCAGCGGCTGTCTGCGACGCC 261
QY 241 GGGGAGGCTTGGCGGCTGCGGCTGTGTCTCTTGGATGAGATGACGGTACTGT 300
DB 262 GGGGAGGACCCGGTGGCGGGGGGCTGTGCTCTTGGAGAGGACGACGACGAGCTGT 321
QY 301 GAGGTGAATGGCCGAGGTACTGTGATGAGAGACCTTTAAACCAATTGACGGTCTG 360
DB 322 GAGGTGAACGGCCGCTGTATCGGGAAGGAGACCTTCCAGCCCACTGACGATCCGC 381
QY 361 TGCCGCTGTATGACGGGTGCTTACCTGCTGCGCTGCTGTGACGTGAGATGTGCGCTG 420
DB 382 TGCCGCTGTGAGAGACGGCGCTTACCTGCTGCGCTGTGACGTGAGATGTGCGCTG 441
QY 421 CCCAGCTGGGACTGCCCCAGCCCCAAGAGATACAGGTGCCAGGAAGTCTGCCCCAG 480
DB 442 CCCAGCTGGGACTGCCCCAGCCCCAAGAGGTGCCAGGTCTGCGGCAAGTCTGCCCCAG 501
QY 481 TGGGTATGTGACCAAGGAGTGCACACCGCGATCCAGCGCTCCAGCGGCCAAGGACCA 540
DB 502 TGGGTGTGGGCCAAGGAGGGGAGCTGGGGAACCCAGCCCTTCCAGCCCCAAGGACCCAG 561
QY 541 CTTTCTGCGCTGTACTCTGCTGCTGTGATGCTCTCTTGTCCAAATTGGAGCACAGCC 600
DB 562 TTTTCTGCGCTGTCTCTTCCCTGCCCCCTGCTGTCCCTGCCCCAGATGAGCACGGCC 621
QY 601 TGGGGCCCTGTCAACCACTGTGGCTGGGATAGCCACCCAGTGTCCAACCAAGAAC 660
DB 622 TGGGACCTGTGCTGACCACTGTGGCTGGGATAGCCACCCGGGTGTCCAACCAAGAAC 681
QY 661 CGATTTCGCAACTGAGATCCAAAGCGGCTGTGTCTGCCAGACCTGCGCTGCGAGCC 720
DB 682 CGCTTGTGCGCACTGAGACCAAGCGCGGCTGTGTCTGTCCAGGGCCCTGCCACCCCTCC 741
QY 721 AGGAGCACAGCTCATGGAACAGTGTCTTA 752
DB 742 AGGGGTGCGAGTCCACAACAAGTGTCTTA 773

RESULT 4
US-09-182-145-14/c

; Sequence 14, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 14
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-14

Query Match 67.8%; Score 510.4; DB 4; Length 1293;
Best Local Similarity 79.9%; Pred. No. 1.4e-122;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGAGCCCACTGATCCATCTTCTGCGCACTTCCCTCTGCTCTCTCAATG 60
DB 1272 ATGAGAGGACACCGAAGACCACCTCTGCGCTTCTCCCTCTGCTCTCTCAAG 1213
QY 61 GTGTGCCCCAGCTGTGCGGACACCTGTACTGTCTGTGAGACCAACCCAGTGGCCA 120
DB 1212 GTGCGTACCAGCTGTGCGCGACACCACTGTACTGTGCGCGCACTTCCCTCTCAAG 1153
QY 121 CAGGGGTACCCCTGTGTCTGATGGCTGTGGCTGCTGTAAGTGTGCACGAGGCTG 180
DB 1152 CTGGAGTACCCCTGTGTCTGATGGCTGTGGCTGCTGCCGGTATGTGCACGGCGCTG 1093
QY 181 GGGAGTCCCTGCGACCACTGCATGTCTGCGACCCCAAGGCCCTGTTGTACGCT 240
DB 1092 GGGAGCCCTGCGACCACTCCACGCTGTGCGACGCCAGCGGCTGTGCCAGCCC 1033
QY 241 GGGGAGGCTTGGCGGCTGCGGCTGTGTCTCTTGGATGAGATGACGGTACTGT 300
DB 1032 GGGGAGGACCCGGTGGCGGGGGGCTGTGCTCTTGGAGAGGACGACAGCAGCTGT 973
QY 301 GAGGTGAATGGCCGAGGTACTGTGATGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
DB 972 GAGGTGAACGGCCGCTGTATCGGGAAGGAGACCTTCCAGCCCCAAGTGTGAGTCCGC 913
QY 361 TGCCGCTGTATGACGGGTGCTTCACTGCTGCGCTGCGCTGTGAGATGAGATGTGCGCTG 420
DB 912 TGCCGCTGTGAGAGACGGCGCTTCACTGCTGCGCTGTGAGAGATGTGCGCTG 853
QY 421 CCCAGCTGGGACTGCCACGCCCAAGAGATACAGGTGCCAGGAAGTGTGCCCGAG 480
DB 852 CCCAGCTGGGACTGCCACGCCCAAGAGGTGCCAGGTCTGCGCAAGTGTGCCCTGAG 793
QY 481 TGGGTATGTGACCAAGGAGTGCACACCGCGATCCAGCGCTCCAGCGGCCAAGACCA 540
DB 792 TGGGTGTGGGCCAAGAGAGGGGAGCTGGGGAACCCAGCCCTTCCAGCCCCAAGACCCAG 733
QY 541 CTTTCTGCGCTGTACTCTGCTGTCTGTGATGCTCTCTTGTCCAAATTGGAGCACAGCC 600
DB 732 TTTTCTGCGCTGTCTCTTCCCTGCCCCCTGTGTGCCCTGCCCCAGATGAGCACGGCC 673

QY	601	TGGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCGAAGC	660
Db	672	TGGGGACCCCTGCTCGACACCACTGTGGGCTGGGCATGCGCACCCGGGTGTCCAACCGAAGC	613
QY	661	CGATTCTGCGCAACTGGAGATCCAAGCGCCCTGTGTGTGTGCCAGACCCCTGGCAGCC	720
Db	612	CGCTTCTGCGCGACTGGAGACCCAGCGCCCTGTGCTGTCCAGGCGCCCTGCCACCCCTCC	553
QY	721	AGGAGCCCAACAGCTCATGGAACAGTGCCTTCTA	752
Db	552	AGGGGTGCGAGTCCACACAAACAGTGCCTTCTA	521

RESULT 5
US-09-182-145-38
; Sequence 38, Application US/09182145B

```

: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hilla, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182,145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 38
: LENGTH: 738
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-182-145-38

```

Query Match	66.6%;	Score 501.2;	DB 4;	Length 738;
Best Local Similarity	79.9%;	Pred. No. 2.9e-120;		
Matches 590; Conservative	0;	Mismatches 148;	Indels 0;	Gaps 0;

OY		13	CCACTGATCCATCTTCTGGCCACCTTCCTCCTCTGTCAATGCTGTGTGCCAG	72
Dd		1	CCGAAGACCACCTCCTGGCCCTTCTTCCTCTCTGCCTCGTCCAAGTGCGTACCAG	60
OY		73	CTGTGCCGGACACCCCTGTACTGTCTGTGGACACCAACCCGAGTGCACAGGGGTACC	132
Dd		61	CTGTGCCCGACACCATGTACTGTGCCCTGGCCACCTCCCGATGCCCGCTGGGAGTACC	120
OY		133	CTGTGCTGGATGGCTGTGGCTGTCTGTAAGAAGTGTGCACAGAGGCTGGGGAGTCTGC	192
Dd		121	CTGTGTCTGGATGGCTGTGGCTGTGCCGGGTATGTGCACAGGCGGCTGGGGAGCCCTGC	180
OY		193	GACCACCTGCATGTCTGCGAACCCACGACGAGGCTGTGTGTACGCTGGGCGACGCCCT	252
Dd		181	GACCAACTCCACGTCTGCGACGCGCACGAGGCGCTGTGTGCCAGCCCGGGGACGACCC	240
OY		253	GCGCGCCATGGGGCTGTGTCTCTTTGATGAGGATGACGGTAGCTGTGAGGTGAATGGC	312
Dd		241	GGTGGCCCCGGGGCCCTGTGCTCTTGGCAGAGGACGACAGCAAGCTGTGAGGTGAACGGC	300
OY		313	CGCAGGTACCTGGATGGAGAGACCTTTAAACCAAATTGCAGGGTCTGTGCCCTGTGAT	372
Dd		301	CGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCCACTGCAGCATCCGCTGCCGCTGCCAG	360

OY	373	GACGGTGCCTTCACCTGCGCTGCCGTGTGCAGTGAGGATGTGCGGCTGCCAGCTGGGAC	432
Db	361	GACGGCGGCTTCACCTGCGCTGCCGTGTGCAGCGAGGATGTGCGGCTGCCAGCTGGGAC	420
OY	433	TGCCCCACGCCCCAAGAATACAGTGTCCAGGAAGTAGTGCTGCCCGAGTGGGTATGTGAC	492
Db	421	TGCCCCACCCAGAGGGGTGAGTGTCTGGCCAAGTGCTGCCCTGAGTGGGTGTGCGGC	480
OY	493	CAGGAGTGACACCGGCGATCCAGCGCTCCACGGCGCAAGGACACCACCTTTCTGCCCTT	552
Db	481	CAAGGAGGGGACTGGGACCCAGCCCCTTCCAGCCCCAAGACCCCACTTTCTGGCCTT	540
OY	553	GTCACCTCCTGCCCTGTGCTGATGCTCCTGTGCCAAATTGGAGCACAGCCTGGGGCCCCCTGC	612
Db	541	GTCCTCTCCCTGCCCCCTGGTGTCCCTGCCCAAGATGGAGCACGGCCTGGGGACCCCTGC	600
OY	613	TCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAACCGATTCTGCCAA	672
Db	601	TGCACCACTGTGGGCTGGGCATAGCCACCCGGGTGTCCAACCAAGAACCGCTTCTGCCGA	660
OY	673	CTGGAGATCCAACGCGGCTGTGTCTGCCCAAGACCCTGCCCTGGCAGCCAGGACCCACAGC	732
Db	661	CTGGAGACCCAGCGCGCTGTGCTGTCCAGGCCCTGCCACACCTCCAGGGGTGCCAGT	720
OY	733	TCATGGAACAGTGCTTTC	750
Db	721	CCACAATAACAGTGCCCTTC	738

RESULT 6
US-09-182-145-39
; Sequence 39, Application US/09182145B

```

: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182,145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 39
: LENGTH: 841
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1-841
: OTHER INFORMATION: Sequence is synthesized.
: Patent No. 6387657
US-09-182-145-39

```

Query Match	65.9%;	Score 496.4;	DB 4;	Length 841;
Best Local Similarity	79.7%;	Pred. No. 5.2e-119;		
Matches 599; Conservative	0;	Mismatches 151;	Indels 2;	Gaps 1;
QY	1	ATGAGGGCAGCCACACTGATCCATCTTCTGGCCACTTCCCTTCTGCGCTTCTCTCAATG	60	

Db 12 ATGAGAGGCACACCGAAGACCACCTCCTGGCCTTCTCCCTCTGCTGCTCTCAAG 71
QY 61 GTGTGTGCCAGCTGTGCCGAGACACCTGTACCTGTCTTGAGACACCCCACTGCCA 120
Db 72 GTGCGTACCACTGTGCCGAGACACCATGTACCTGCTCCCTGCGACCTCCCGATGCCG 131
QY 121 CAGGGGTACCCCTGTGTCTGTGATGGCTGTGGCTGTAAAGTGTGTGACGAGGCTG 180
Db 132 CTGGAGTACCCCTGTGTGTGATGGCTGTGGCTGTCCGGGTATGTGACGCGGCTG 191
QY 181 GGGAGTCTGTGACCACTGCATGTCTGCGACCCAGGCGCTGTGTGACCT 240
Db 192 GGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGGCGCTGTGTGACGCTG 251
QY 241 GGGGAGCCCTGGCGCCATGGGCTGTGTCTGTGATGAGATGAGGCTGCTGT 300
Db 252 GGGGAGGACCCGCTGTGCGGGGGGGGCGCTGTGTGCGAGAGGACGACGCTGT 311
QY 301 GAGGTGAATGCCCGAGGTACCTGTGATGAGAGACCTTAAACCAATGCGAGGCTCTG 360
Db 312 GAGGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCGACATGACATCCGC 371
QY 361 TGCCGCTGTATGAGCGTGGCTTACCTGCTGCTGCTGTGACGTGAGATGTGCGCTG 420
Db 372 TGCCGCTGCGAGAGCGCGCTTACCTGCTGCTGCTGTGACGAGATGTGCGCTG 431
QY 421 CCCAGCTGGAGTGGCCACGCCCCAAGAGATACAGGTGCCAGGAAGTGTGCCCCGAG 480
Db 432 CCCAGCTGGAGTGGCCACGCCCCAAGAGAGGTGCCAGGTGCCAGGTGTGCCCCGAG 491
QY 481 TGGGTATGTACCAAGGAGTACACCGCGGCTGACGCGCTCCAGCGCGCAAGAGACCAA 540
Db 492 TGGGTGTGCGGCGCAAGAGGGGAGTGGGAGACGCGCTTCCA--GCCAAGAGACCCAG 549
QY 541 CTCTTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 600
Db 550 TTTCTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 609
QY 601 TGGGCGCCCTGTCAACCACTGTGGGCTGGGATGACCAACCGAGTGTCCAACAGAAC 660
Db 610 TGGGAGCCCTGTCAACCACTGTGGGCTGGGATGACCAACCGAGTGTCCAACAGAAC 669
QY 661 CGATTCTGCAACTGAGATCAACGCGCGCTGTCTGTGCGCAGACCTGCGTGGCAGCC 720
Db 670 CGCTTCTGCGAGTGAAGACCAAGCGCGCTGTGCTGTGCGCAGCGCTGCGCAGCCCTCC 729
QY 721 AGGAGCCACAGCTCATGGAACAGTGTCTCA 752
Db 730 AGGGGTGCGAGTCCACAAACAGTGTCTCA 761

RESULT 7
US-08-167-628-1
; Sequence 1, Application US/08167628
; Patent No. 5408040
; GENERAL INFORMATION:
; APPLICANT: Grotenhorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Judas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
US-08-167-628-1

Query Match 21.6%; Score 163; DB 1; Length 2075;
Best Local Similarity 56.2%; Pred. No. 4.5e-33;
Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

QY 102 GACACACACCCAGTGGCCACAGGGGTACCCCTGTGCTGTGATGGCTGTGCTGTAA 161
Db 243 GCCGGCGCGCGCTGCGCGCGCGCGCGCGCGCTGAGCTGTGCTGAGCGCTGCTGCCG 302
QY 162 AGTGTGTGACGAGGCTGGGGAGTCTGTGCGACCACTGCATGTGTGCAACCCAGCCA 221
Db 303 CGTCTGCGCAAGACAGCTGGGCGAGCTGTGACCGAGCGGACCCCTGCGACCCGACAA 362
QY 222 GGGCTGTGTGTGACCTGGGCGAGGCGCTGCGCGCGCGCGCGCGCTGTGTCTTGA 281
Db 363 GGGCTCTTGTGTGACTTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 419
QY 282 TGAAGATGACGCTGAGTGTGAGTGAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 341
Db 420 CAAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
QY 342 ACCCAATGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 401
Db 480 GAGCAGCTGCAAGTACCAAGTGCACGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 539
QY 402 CAGTGAAGATGTGGCGCTGCGCGAGCTGGAGTGGCCACGCGCGCGCGCGCGCGCGCG 461
Db 540 CAGCATGAGCTGCTGCTGCTGCGCGAGCGCTGAGTGGCGCGCGCGCGCGCGCGCGCG 599
QY 462 AGGAAGTGTGCGCGCGAGTGGGTATGTGACCAAGGAGTGCACACCGCGCGCGCGCG 521
Db 600 CGGGAATGTGCGCGAGAGTGGGTGTGTGACGAGCGCGCGCGCGCGCGCGCGCGCG 659
QY 522 CAGGCGCGCAAGGACCAACTTCTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 575
Db 660 TGCCCTGCGCGCGCTTACCGAGCTGGAAGACACGTTGGCGCGCGCGCGCGCGCGCG 719
QY 576 TCCTTGTCAAAATGAGACAGCAGCGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 635
Db 720 CAACTGCTGTGTCAGACACAGAGTGAAGCGCGCTGTCCAAAGACCTGTGGATGGCAT 779
QY 636 AGCCACCCGAGTGTCAACAGAGACCGAGTGTGCAACTGAGATGCAACCGCGCGCTGTG 695
Db 780 CTCACACCGGTTACCAATGACACAGCGCTCTGCAAGGCTAGAGAAGCAGAGCGCGCTGTG 839
QY 696 TCTGCCAGACCTGCGCTGGCAGCC 720

```

Db      840 CATGCTCAGGCGCTTGCGAAGCTGAC 864

RESULT 8
US-08-386-680-1
; Sequence 1, Application US/08386680
; Patent No. 5585270
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/386,680
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
; US-08-386-680-1

Query Match      21.6%; Score 163; DB 1; Length 2075;
Best Local Similarity 56.2%; Pred. No. 4.5e-33;
Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2

QY      102 GACACCACCCAGTGCACAGGGGGGTACCCCTGTGCTGATGAGCTGTGGCTGTGTA 161
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      243 GCCGGCGCGCGCTGCCCGCGGGGGCTGAGCCTGTGCTGAGCGCTGCGGCTGCGCG 302
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      162 AGTGTGTGACGAGAGCTGGGGAGTCTCTGCGACCACTGATGCTGCGACCCAGCCA 221
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      303 CGTCTGCGCAAGCAGCTGGGCGAGCTGTGCACCGAGCGCACCCCTGCGACCCGACAA 362
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      222 GGGCCTGTTGTACGCTGGGGCAGGCCCTGGCGGCATGGGCTGTGTCTTGA 281
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      363 GGGCCTCTTCTGTGACTTCGGCTCCCGGCCAACCGCAAGTTCGGCGTGTG---CACCGC 419
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      282 TGAAGATGACGATAGCTGTGAGCTGAATGGCCGCGAGGTACCTGATGATGAGAGACCTTAA 341
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      420 CAAAGATGTGTCTCCCTGCATCTTCGGGTGTGACGGTGTACCGCAGCGGAGAGTCTTCCA 479
      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

QY 342 ACCCAATTGCAGGGTCTGTCCCGCTGTGATGACGGGTGGCTTCACCCTGCCCGCTGTG 401
|| ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 480 GAGCAGCTGCAAGTACCAGTGACGTGCTGGACGGGGGGGTGGGCTGCATGCCCTGTG 539

QY 402 CAGTAGGATGTGGCGCTGCCAGCTGGGACTGCCACCAGCCCCAAGAATAACAGGTGCC 461
|| | || | || | ||| ||| | ||| ||| | || | || | |||
Db 540 CAGCATGGACGTTGTTGTGCCCCAGCCCTGACTGCCCTTCCCAGAGGAGGTCAGCTGCC 599

QY 462 AGGAAAGTGTGCCCCGAGTGGGTATGTACCAAGGAGTGAACCGCGATCCAGCGCTC 521
|| | ||| ||| | ||| ||| ||| | || | | | | | | |
Db 600 CGGGAATGCTGCGAGAGTGGGTGTGTGACGAGCCCAAGAACCAACCGGTGTGGGCC 659

QY 522 CACGGCGCAAGGACACCAACTTTTGCCCTTGTCACTCCTGCCTGTGAT-----GC 575
| | | | ||| ||| | | | | | | | | | | | |
Db 660 TGCCCTCGCGCTTACCGACTGGAAGACAGTTTGGCCAGACCCAACTAATATTAGAGC 719

QY 576 TCCTTGTCAAATTTGAGCACAGCCTGGGGCCOCTGCTCAACCACCTGTGGGCTGGGCAT 635
|| | || | ||| ||| ||| ||| ||| | | | ||| ||| |||
Db 720 CAACTGCCCTGTGCCAGACCACAGAGTGGAGCGCCTGTTCAGAGACCTGTGGGATGGGCAT 779

QY 636 AGCCACCCGAGTGTCCAACCAAGAACCGATTCTGCCAAGTGGAGATCCAACGCCCGCTGTG 695
||||| || | ||| | ||| | ||| | ||| ||| ||| ||| |||
Db 780 CTCACACCCGGGTTACCAATGACAACGCCCTCTGCAGGCTAGAGAAGCAGAGCGCCTGTG 839

QY 696 TCTGCCCAGACCCCTGCCTGGCAGCC 720
|| | ||| ||| ||| | | |
Db 840 CATGTCAGGCCCTTGCGAAGCTGAC 864

```

RESULT 9
US-08-459-717-1
; Sequence 1, Application US/08459717
; Patent No. 5770209
;
; GENERAL INFORMATION:
;
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubb & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,717
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,427
; FILING DATE: 30-AUG-1991
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
;

```


Db 720 CAACTGCTGTGTCAGACACAGAGTGAGCGCTGTTCGAGACCTGTGGATGGGCAT 779
QY 636 AGCCACCCGAGTGTCCACCAACCGATTTCTGCCACTGGAGATCCACGCCGCTGTG 695
Db 780 CTCACCCCGGGTTACCAATGACACAGCCCTCCTGACAGGCTAGAGAAGCAGAGCCGCTGTG 839
QY 696 TCTGCCACAGACCCTGCTGGCAGCC 720
Db 840 CATGCTCAGGCTTGGCAGCTGAC 864

RESULT 11

US-08-880-031-1
; Sequence 1, Application US/08880031
; Patent No. 5916756
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,031
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: DB60R32
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1177
; US-08-880-031-1

Query Match 21.6%; Score 163; DB 2; Length 2075;
Best Local Similarity 56.2%; Pred. No. 4.5e-33;
Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

QY 102 GACACCAACCCAGTGCACACAGGGGTACCCCTGTGCTGATGGCTGTGGCTGTAA 161
Db 243 GCCGGCGCGCGCTGCCCGGGCGGCTGAGCCTGCTGGACGCTGGCGCTGTGCCG 302
QY 162 AGTGTGTCACGAGGCTGGGGAGTCTGCGACCACTGCATGCTGCGACCCAGCCA 221
Db 303 CGTCTGGCCCAAGCAGCTGGGCGAGCTGTGCACCGAGCGGACCCCTGCGACCCGACAA 362
QY 222 GGGCTGTGTTGTACGCTGGGGCAGCCCTGGCGGCATGGGGCTGTGTCTTGA 281

Db 363 GGGCTCTTCTGTGACTTGGGCTCCCGGCCAACCCGCAAGATCGCGTGTG---CACCC 419
QY 282 TGAGATGACGGTAGCTGTGAGGTGAATGCGCCGAGGTACTGTGATGAGAGACTTTAA 341
Db 420 CAAGATGTGCTCCCTGCATCTTCGGTGTAGCGGTACCGCAGCGAGAGTCTTCCA 479
QY 342 ACCCAATTGACGGGTCTGTGCCGCTGTGATGACGGGTGCTTCACTGCTGCCCTGTG 401
Db 480 GAGCAGCTGCAAGTACCAAGTGCACAGCTGCGCTGAGCGGGGGGTGGCTGCATGCCCTGTG 539
QY 402 CAGTGAAGATGTGCGGCTGCCAGCTGGGACTGAGCCACGCCCAAGAGATACAGTGCC 461
Db 540 CAGCATGAGCTGTGCTGTGCCAGCCCTGACTGCCCTTCCGAGAGGGTCAAGCTGCC 599
QY 462 AGGAAGTGTGCTGCCCGAGTGGGTATGTGACCAAGGAGTGCACACCGGATCCAGCGCTC 521
Db 600 CGGGAATGCTGCGAGAGTGGGTGTGTGACGAGGCCCAAGACCAACCGTGGTGGGCC 659
QY 522 CAGGCGCAAGACACCACTTTCTGCCCTTGTACATCTCTGCTGTGAT-----GC 575
Db 660 TGCCCTCGCGGCTTACCGACTGGAAGACACAGTGTGGCCAGACCACTATGATTAGAGC 719
QY 576 TCTGTGTCAAATTGAGACAGCCTGGGGCCCTGTCAACCACTGTGGGCTGGGCAT 635
Db 720 CAACTGCTGTGTCAGACACAGAGTGGAGCGCCTGTTCAGAAGACCTGTGGATGGGCAT 779
QY 636 AGCCACCCGAGTGTCCAACCAACCAACCGATTTCTGCCAAGTGAATCCAAAGCGGCTGTG 695
Db 780 CTCACCCCGGGTTACCAATGACACAGCCTCCTGACAGGCTTAGAGAAGCAGAGCCGCTGTG 839
QY 696 TCTGCCACAGACCCTGCTGGCAGCC 720
Db 840 CATGCTCAGGCTTGGCAGCTGAC 864

RESULT 12

US-09-097-179-1
; Sequence 1, Application US/09097179
; Patent No. 6149916
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,179
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,680
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: CTGF
FEATURE:
NAME/KEY: CDS
LOCATION: 130..1177
PCT-US96-08140-1

Query Match 21.6%; Score 163; DB 5; Length 2075;
Best Local Similarity 56.2%; Pred.No.4.5e-33;
Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

QY 102 GACACCACCCAGTGGCCACAGGGGTACCCCTGGTGTGATGGCTGTGCTGTAA 161
Db 243 GCCGGCGCCGCGCTGCCCGCGCGGTGAGCCTCGTGTGAGCGGCTGGGCTGCGCG 302
QY 162 AGTGTGTGACGAGGCTGGGGAGTCTCGACCACTGCATGTGCGACCCAGCCA 221
Db 303 CGTCTGCGCCAGACGAGCTGGGGCAGCTGTGACCCGAGCGGACCCCTGCGACCCGACAA 362
QY 222 GGGCCCTGCTTGTACGCTGGGGCAGGGCCCTGGCGGCCCATGGGCTGTGTCTCTTGA 281
Db 363 GGGCCTCTTCTGTACTTCGGCTCCCGGCCAACCGCAAGATCGCGGTGTG--CACCGC 419
QY 282 TGAGGATGACGGTAGCTGTGAGGTGAATGGCCGAGTACCTGGATGAGAGACCTTAA 341
Db 420 CAAGATGTGCTCCCTGCATCTTCGGTGTAGCGGTACCGCAGCGAGAGTCTTCCA 479
QY 342 ACCCAATTGCAGGGTCCCTGCGCGCTGTGATGACGGGTGCTTCACTGCGCGCTGTG 401
Db 480 GAGCAGCTGCAAGTACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 539
QY 402 CAGTGAGATGTGCGGCTGCCAGCTGGGACTGCCACGCCCAAGAGATACAGGTGCC 461
Db 540 CAGCATGAGACGTCTGCTGCCACCCCTGACTGCCCTTCCGAGAGGCTCAAGCTGCC 599
QY 462 AGGAAGTGTGCTGCCGAGTGGTATGTGACCAAGGAGTACACCGCGATCCAGCGCTC 521
Db 600 CGGGAATGCTGCGAGAGTGGGTGTGTACGAGGCCCAAGAGACCAACCGTGTGGGCC 659
QY 522 CACGGCGCAAGACCACTTCTGCCCTGTCACTCTGCTCTGCTGAT-----GC 575
Db 660 TGCCCTCGCGGCTTACCGACTGGAAGACACGTTTGGCCAGACCCCAACTATGATTAGAGC 719
QY 576 TCCTGTCCAATTTGAGACAGCTGGGGCCCTGCTCAACCACTGTGGGCTGGCAT 635
Db 720 CACTGCTGTGTCCAGACACAGAGTGGAGCGCTGTCCAAGACCTGTGGATGGCAT 779
QY 636 AGCCACCCGAGTGTCCAACCAAGACGATTTGCAACTGAGATCCAAGCGCGCTGTG 695
Db 780 CTCCACCCGGGTATCAATGACACAGGCTCTCTGACAGGCTAGAGAGACAGCGCGCTGTG 839
QY 696 TCTGCCAGACCTGCTGGCAGCC 720
Db 840 CATGTCAGGCTTGGCAAGCTGAC 864

Search completed: July 28, 2003, 21:36:01
Job time : 43.1673 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:36:41 ; Search time 185.622 Seconds
(without alignments)
8368.853 Million cell updates/sec

Title: US-10-010-408-3
Perfect score: 753
Sequence: 1 ATGAGGGGCGACCCACTGAT.....CATGGAACAGTGCTTCTTAA 753

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCRUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	753	100.0	753	15	US-10-010-408-3
2	753	100.0	1708	15	US-10-010-408-1
3	681	90.4	681	15	US-10-010-408-12
4	659	87.5	1734	15	US-10-112-267-17
5	659	87.5	1734	15	US-10-112-267-18
6	510.4	67.8	1266	14	US-10-137-866-319
7	510.4	67.8	1266	14	US-10-146-726-319
8	510.4	67.8	1266	14	US-10-146-727-319
9	510.4	67.8	1266	14	US-10-146-788-319
10	510.4	67.8	1266	14	US-10-152-380-319
11	510.4	67.8	1266	15	US-10-028-072-319
12	510.4	67.8	1266	15	US-10-121-049-319
13	510.4	67.8	1266	15	US-10-123-904-319
14	510.4	67.8	1266	15	US-10-140-470-319
15	510.4	67.8	1266	15	US-10-175-746-319
16	510.4	67.8	1266	15	US-10-175-746-319

17	510.4	67.8	1266	15	US-10-176-918-319	Sequence 319, App
18	510.4	67.8	1266	15	US-10-176-921-319	Sequence 319, App
19	510.4	67.8	1266	15	US-10-137-865-319	Sequence 319, App
20	510.4	67.8	1266	15	US-10-140-474-319	Sequence 319, App
21	510.4	67.8	1266	15	US-10-142-431-319	Sequence 319, App
22	510.4	67.8	1266	15	US-10-143-114-319	Sequence 319, App
23	510.4	67.8	1266	15	US-10-140-002-319	Sequence 319, App
24	510.4	67.8	1266	15	US-10-142-419-319	Sequence 319, App
25	510.4	67.8	1266	15	US-10-123-262-319	Sequence 319, App
26	510.4	67.8	1266	15	US-10-142-423-319	Sequence 319, App
27	510.4	67.8	1266	15	US-10-121-050-319	Sequence 319, App
28	510.4	67.8	1266	15	US-10-141-755-319	Sequence 319, App
29	510.4	67.8	1266	15	US-10-143-032-319	Sequence 319, App
30	510.4	67.8	1266	15	US-10-123-108-319	Sequence 319, App
31	510.4	67.8	1266	15	US-10-123-236-319	Sequence 319, App
32	510.4	67.8	1266	15	US-10-123-261-319	Sequence 319, App
33	510.4	67.8	1266	15	US-10-140-921-319	Sequence 319, App
34	510.4	67.8	1266	15	US-10-140-928-319	Sequence 319, App
35	510.4	67.8	1266	15	US-10-121-045-319	Sequence 319, App
36	510.4	67.8	1266	15	US-10-123-292-319	Sequence 319, App
37	510.4	67.8	1266	15	US-10-123-903-319	Sequence 319, App
38	510.4	67.8	1266	15	US-10-124-819-319	Sequence 319, App
39	510.4	67.8	1266	15	US-10-140-925-319	Sequence 319, App
40	510.4	67.8	1266	15	US-10-160-498-319	Sequence 319, App
41	510.4	67.8	1266	15	US-10-124-824-319	Sequence 319, App
42	510.4	67.8	1266	15	US-10-127-825A-319	Sequence 319, App
43	510.4	67.8	1266	15	US-10-127-829A-319	Sequence 319, App
44	510.4	67.8	1266	15	US-10-127-835A-319	Sequence 319, App
45	510.4	67.8	1266	15	US-10-127-835A-319	Sequence 319, App

ALIGNMENTS

RESULT 1
US-10-010-408-3
Sequence 3, Application US/10010408
Publication No. US20020165185A1
GENERAL INFORMATION:
APPLICANT: John J. Castelliott, Jr.
TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules and Uses Therefor
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 753 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..750
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-010-408-3

Query Match 100.0%; Score 753; DB 15; Length 753;
Best Local Similarity 100.0%; Pred. No. 7.4e-206;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGAGGGGAGCCGACCTGATCCATCTTGTGGCCACTTCTCTCTGCTGCTCTCTCAATG 60
Db 1 ATGAGGGGAGCCGACCTGATCCATCTTGTGGCCACTTCTCTCTGCTGCTCTCTCAATG 60
OY GTGTGTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGGCCA 120
Db GTGTGTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGGCCA 120
OY 121 CAGGGGGTACCCCTGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACAGGAGCTG 180
Db 121 CAGGGGGTACCCCTGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACAGGAGCTG 180
OY 121 CAGGGGGTACCCCTGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACAGGAGCTG 180
Db 121 CAGGGGGTACCCCTGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACAGGAGCTG 180
OY 181 GGGGAGTCTCGGACACCTGATGTCTGCGACCCAGCGAGGGCTGTGTGACGCT 240
Db 181 GGGGAGTCTCGGACACCTGATGTCTGCGACCCAGCGAGGGCTGTGTGACGCT 240
OY 241 GGGGAGGAGGGCTGGCGGCGGCTGTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 300
Db 241 GGGGAGGAGGGCTGGCGGCGGCTGTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 300
OY 301 GAGGTGAATGGCGCAGGTGATGAGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
Db 301 GAGGTGAATGGCGCAGGTGATGAGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
OY 361 TGCCGCTGTGATGACGGTGTCTACCTGCGCTGCGCTGCGCTGCGCTGCGCTG 420
Db 361 TGCCGCTGTGATGACGGTGTCTACCTGCGCTGCGCTGCGCTGCGCTGCGCTG 420
OY 421 CCCAGCTGGGACTGCCACGCCCCAAGAATACAGGTGGCAGGAAAGTGTGCCCCGAG 480
Db 421 CCCAGCTGGGACTGCCACGCCCCAAGAATACAGGTGGCAGGAAAGTGTGCCCCGAG 480
OY 481 TGGGTATGTGACAGGAGTGAACACCGCGGATCCAGCGCTCCACGCGGCAAGACCAA 540
Db 481 TGGGTATGTGACAGGAGTGAACACCGCGGATCCAGCGCTCCACGCGGCAAGACCAA 540
OY 541 CTCTTCTGCCCTGTGCTACTCTCTGCTCTGCTGATGCTCTCTTCCAAATTGGAGCACAGCC 600
Db 541 CTCTTCTGCCCTGTGCTACTCTCTGCTCTGCTGATGCTCTCTTCCAAATTGGAGCACAGCC 600
OY 601 TGGGGCCCCCTGCTCAACACCTGTGGGCTGGGATAGCCAGCCGAGTGTCCAAACAGAAC 660
Db 601 TGGGGCCCCCTGCTCAACACCTGTGGGCTGGGATAGCCAGCCGAGTGTCCAAACAGAAC 660
OY 661 CGATTCTGCAACTGAGATCCAAACGCGGCTGTGTGCGCAGACCCCTGCGGAGGCC 720
Db 661 CGATTCTGCAACTGAGATCCAAACGCGGCTGTGTGCGCAGACCCCTGCGGAGGCC 720
OY 721 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 753
Db 721 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 753

RESULT 2
US-10-010-408-1
; Sequence 1, Application US/10010408
; Publication No. US20020165185A1
; GENERAL INFORMATION:

APPLICANT: John J. Castelli, Jr.
TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CCN-Like Molecules
and Uses Therefor
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandiagouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1708 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 249..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-010-408-1

Query Match 100.0%; Score 753; DB 15; Length 1708;
Best Local Similarity 100.0%; Pred. No. 8.2e-206;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGAGGGGAGCCGACCTGATCCATCTTGTGGCCACTTCTCTCTGCTGCTCTCTCAATG 60
Db 249 ATGAGGGGAGCCGACCTGATCCATCTTGTGGCCACTTCTCTCTGCTGCTCTCTCAATG 308
OY 61 GTGTGTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGGCCA 120
Db 309 GTGTGTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGGCCA 368
OY 121 CAGGGGGTACCCCTGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACAGGAGCTG 180
Db 369 CAGGGGGTACCCCTGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACAGGAGCTG 428
OY 181 GGGGAGTCTCGACACCTGATGTCTGCGACCCAGCGAGGGCTGTGTGACGCT 240
Db 429 GGGGAGTCTCGACACCTGATGTCTGCGACCCAGCGAGGGCTGTGTGACGCT 488
OY 241 GGGGAGGGCTGTGCGGCGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 300
Db 489 GGGGAGGGCTGTGCGGCGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTGT 548
OY 301 GAGGTGAATGGCGCAGGTACCTGATGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
Db 549 GAGGTGAATGGCGCAGGTACCTGATGAGAGACCTTTAAACCAATTGACAGGCTCTG 608

```

QY 361 TGCCGCTGTGATGACGGTGGCTTCACTGCGCTGCCCTGTGACGTAGAGATGTGGGCTG 420
Db 609 TGCCGCTGTGATGACGGTGGCTTCACTGCGCTGCCCTGTGACGTAGAGATGTGGGCTG 668
QY 421 CCCAGCTGGAGTCCACGCCCCAAGAGATACAGTGCAGGAAGTGTGCCCCGAG 480
Db 669 CCCAGCTGGAGTCCACGCCCCAAGAGATACAGTGCAGGAAGTGTGCCCCGAG 728
QY 481 TGGGTATGTAGCAGGAGTGAACACCGCGATCCAGCGCTCCACGCGCAAGAGACCA 540
Db 729 TGGGTATGTAGCAGGAGTGAACACCGCGATCCAGCGCTCCACGCGCAAGAGACCA 788
QY 541 CTTTCTGCCCTGTGACCTGCTGCTGCTGATGCTCTGCTGCTCAATTTGAGACAGCC 600
Db 789 CTTTCTGCCCTGTGACCTGCTGCTGCTGATGCTCTGCTGCTCAATTTGAGACAGCC 848
QY 601 TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACAGAAC 660
Db 849 TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACAGAAC 908
QY 661 CGATTCTGCCAAGTGGAGATCCAAACGCGCGCTGTGCTGCCCCAGACCCCTGCGTGGCAGCC 720
Db 909 CGATTCTGCCAAGTGGAGATCCAAACGCGCGCTGTGCTGCCCCAGACCCCTGCGTGGCAGCC 968
QY 721 AGAGCCACAGCTCATGGAACAGTGTCTTAA 753
Db 969 AGAGCCACAGCTCATGGAACAGTGTCTTAA 1001

```

RESULT 3

US-10-010-408-12

; Sequence 12, Application US/10010408
; Publication No. US20020165185A1

; GENERAL INFORMATION:

; APPLICANT: John J. Castelli, Jr.
; TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules
; and Uses Therefor

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston

; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/010,408
; FILING DATE: 07-Dec-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/044,273
; FILING DATE: March 19, 1998

; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..681
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
us-10-010-408-12

```

Query Match 90.4%; Score 681; DB 15; Length 681;
Best Local Similarity 100.0%; Pred. No. 3e-185;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 70 CAGCTGTGCCGACACACCTGTACCTGTCTCTTGAGACACACCCAGTGGCCACAGGGGTA 129
Db 1 CAGCTGTGCCGACACACCTGTACCTGTCTCTTGAGACACACCCAGTGGCCACAGGGGTA 60
QY 130 CCCCTGTGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGAGAGCTGGGGAGTCC 189
Db 61 CCCCTGTGCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGAGAGCTGGGGAGTCC 120
QY 190 TGGCACCACCTGCATGTCTGCGAACCCAGCCAGGGGCTGGTTGTCAAGCTGGGGCAGGC 249
Db 121 TGGCACCACCTGCATGTCTGCGAACCCAGCCAGGGGCTGGTTGTCAAGCTGGGGCAGGC 180
QY 250 CCTGCGGGCCATGGGGCTGTGTCTCTCTTGATGAGATGACGGTAGCTGTGAGTGAAT 309
Db 181 CCTGCGGGCCATGGGGCTGTGTCTCTCTTGATGAGATGACGGTAGCTGTGAGTGAAT 240
QY 310 GGCCGACAGTACCTGATGAGAGAACCTTTAAACCAATTGACAGGCTCCTGTGCGCTGT 369
Db 241 GGCCGACAGTACCTGATGAGAGAACCTTTAAACCAATTGACAGGCTCCTGTGCGCTGT 300
QY 370 GATGACGGTGGCTTACCTGCTGCCCTGCGCTGTGCAAGTGAAGATGTGGGCTGCCAGCTGG 429
Db 301 GATGACGGTGGCTTACCTGCTGCCCTGCGCTGTGCAAGTGAAGATGTGGGCTGCCAGCTGG 360
QY 430 GACTGCCCCACGCCCCAAGAGATACAGGTTGCCAGGAAGTGTGCCCCAGTGGGTATGT 489
Db 361 GACTGCCCCACGCCCCAAGAGATACAGGTTGCCAGGAAGTGTGCCCCAGTGGGTATGT 420
QY 490 GACCAGGAGTGAACACCGCGGATCCAGCGCTCCACGGCGCAAGACACCAACTTCTGCC 549
Db 421 GACCAGGAGTGAACACCGCGGATCCAGCGCTCCACGGCGCAAGACACCAACTTCTGCC 480
QY 550 CTGTCACTCCTGCTGCTGTGATGCTCTTGTCCAAATTGAGACACAGCCTGGGCGCCC 609
Db 481 CTGTCACTCCTGCTGCTGTGATGCTCTTGTCCAAATTGAGACACAGCCTGGGCGCCC 540
QY 610 TGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACAGAACCGATCTGC 669
Db 541 TGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACAGAACCGATCTGC 600
QY 670 CAACTGGAGATCCAAACGCGCTGTGTCTGCCAGAGACCTGCTGGCAGCCAGGAGCCAC 729
Db 601 CAACTGGAGATCCAAACGCGCTGTGTCTGCCAGAGACCTGCTGGCAGCCAGGAGCCAC 660
QY 730 AGCTCATGGAACAGTGTTC 750
Db 661 AGCTCATGGAACAGTGTTC 681

```

RESULT 4

US-10-112-267-17

; Sequence 17, Application US/10112267
; Publication No. US20030068678A1

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-112-267-17

Query Match 87.5%; Score 659; DB 15; Length 1734;
Best Local Similarity 93.0%; Pred. No. 6.6e-179;
Matches 702; Conservative 0; Mismatches 50; Indels 3; Gaps 1;

QY 1 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCAATG 60
Db 257 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCAATG 316
QY 61 GTGTGTGCCAGCTGTGCCGGACACCCCTGTACTGTCTTGGACACACCCAGTGGCCA 120
Db 317 GTGTATTCACAGCTGTGCCAGACACCCCTGTCTGTCTTGGACACACCCAGTGGCCA 376
QY 121 CAGGGGTACCCCTGTGTCTGTGATGGCTGTGCTGTAAAGTGTGTGACGAGGCTG 180
Db 377 CCGGGGTACCCCTGTGTCTGTGATGGCTGTGCTGTGAGTGTGTGACGAGGCTG 436
QY 181 GGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGCCCTGTTGTCAAGCT 240
Db 437 GGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGCCCTGTTGTCAAGCT 496
QY 241 GGGGACAGCCCTGGGGCCATGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT 300
Db 497 GGGGACAGCCCTGGGGCCATGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT 556
QY 301 GAGTGAATGGCCGACAGTACCTGGATGGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
Db 557 GAGTGAATGGCCGACAGTACCTGGATGGGGAGACCTTTAAACCAATTGACAGGCTTTTG 616
QY 361 TGCCGCTGTGATGACGCTGGCTTACCTGCGCTGCTGTGACAGTGAAGATGCGGCTG 420
Db 617 TGCCGCTGTGATGACGCTGGCTTACCTGCGCTGCTGTGACAGTGAAGATGCGGCTG 676
QY 421 CCCAGCTGGGACTGCCACCGCCCAAGAGAATACAGGTGCCAGGAAGTGTGCCCGAG 480
Db 677 CCCAGCTGGGACTGCCACCGCCCAAGAGAATACAGGTGCCAGGAAGTGTGCCCGAG 736
QY 481 TGGGTATGTACCAAGGAGTGA--CACCGGCAATCCAGCGCTCCACGGCGCAAGGACAC 537
Db 737 TGGGTATGTACCAAGGAGTGA--CACCGGCAATCCAGCGCTCCACGGCGCAAGGACAC 796
QY 538 CAACCTTCTGCGCTGTCACTCTGCTGTGCTGTGCTGTCTTGAATTTGAGACACA 597
Db 797 CAACCTTCTGCGCTGTCACTCTGCTGTGCTGTGCTGTCTTGAATTTGAGACACA 856
QY 598 GCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACAG 657
Db 857 GCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACAG 916
QY 658 AACCGATTCTGCCAATGAGATCCAAACGCGCGCTGTGTCTGCCACAGACCTGCGTGA 717
Db 917 AACCGATTCTGCCAATGAGATCCAAACGCGCGCTGTGTCTGCCACAGACCTGCGTGA 976
QY 718 GCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTA 752

Db 977 TCCAGAGCCACGCGCTCATGGAACAGTGCCTTCTA 1011

RESULT 5
US-10-112-267-18/c
; Sequence 18, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 18
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-112-267-18

Query Match 87.5%; Score 659; DB 15; Length 1734;
Best Local Similarity 93.0%; Pred. No. 6.6e-179;
Matches 702; Conservative 0; Mismatches 50; Indels 3; Gaps 1;

QY 1 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCAATG 60
Db 1478 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCAATG 1419
QY 61 GTGTGTGCCAGCTGTGCCGGACACCCCTGTACTGTCTTGGACACACCCAGTGGCCA 120
Db 1418 GTGTATTCACAGCTGTGCCAGACACCCCTGTGTCTTGGACACACCCAGTGGCCA 1359
QY 121 CAGGGGTACCCCTGTGTCTGTGATGGCTGTGCTGTAAAGTGTGTGACGAGGCTG 180
Db 1358 CCGGGGTACCCCTGTGTCTGTGATGGCTGTGCTGTGAGTGTGTGACGAGGCTG 1299
QY 181 GGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGCCCTGTTGTCAAGCT 240
Db 1298 GGGAGTCTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGCCCTGTTGTCAAGCT 1239
QY 241 GGGGACAGCCCTGGGGCCATGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT 300
Db 1238 GGGGACAGCCCTGGGGCCATGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT 1179
QY 301 GAGTGAATGGCCGACAGTACCTGGATGGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
Db 1178 GAGTGAATGGCCGACAGTACCTGGATGGGGAGACCTTTAAACCAATTGACAGGCTTTTG 1119
QY 361 TGCCGCTGTGATGACGCTGGCTTACCTGCTGCTGTGCTGTGACAGTGAAGATGCGGCTG 420
Db 1118 TGCCGCTGTGATGACGCTGGCTTACCTGCTGCTGTGCTGTGACAGTGAAGATGCGGCTG 1059
QY 421 CCCAGCTGGAGTGGCCACGCCCCAAGAGAATACAGGTGCCAGAAAGTGTGCCCCGAG 480
Db 1058 CCCAGCTGGAGTGGCCACGCCCCAAGAGAATACAGGTGCCAGAAAGTGTGCCCCGAG 999

OY 481 TGGTATGTGACCAGGAGTGA---CACCGCGATCCAGCGCTCCACGGCGCAAGACAC 537
||||| ||||||| ||||| ||||||| ||||| ||||||| ||||| ||||||| |||||
Db 998 TGGGTGTGTGACCAGCAGCAGTGTGACGCCGCAATCCAGCCCTCCTCAGCCCAAGACAC 939
OY 538 CAACTTCTGCCCTGTGTACTCCCTGCTCTGCTGATGCTCCTGTGTCCTCAATTGGAGCACA 597
||||| ||||||| ||||||| ||||||| ||||| ||||||| ||||| ||||||| |||||
Db 938 CAACTTCTGCCCTGTGTACTCCTGTGATCTGCCGATGGCCCCCTGTCCAACTGGAGCACA 879
OY 598 GCCTGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAAACCAG 657
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db 878 GCCTGGGCCCCCTGCTCAACCACCTGTGGGTTGGGCATAGCCACCCGAGTATCCAAACCAG 819
OY 658 AACCGATTCTGCCAAGTGAAGATCCAGCCGCCCTGTGTCTGCCAGACCCCTGCCTGCA 717
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db 818 AACCGATTCTGCCAAGTGAAGATCCAGCGCTGCCTGTGTCTGTCTCCAGACCCCTGCCTGCA 759
OY 718 GCCAGAGCCACAGCTCATGAACAGTGTCTTCTA 752
||||| ||||||| ||||||| ||||||| |||||
Db 758 TCCAGAGCCACAGGCTCATGAACAGTGTCTTCTA 724

RESULT 6
US-10-137-866-319: Sequence 319, Application US/10137866
: Publication No. US20030129689A1

: GENERAL INFORMATION:

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerltsen, Mary E.
: APPLICANT: Goddard, Audrey J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3330R1C151
: CURRENT APPLICATION NUMBER: US/10/137, 866
: PRIOR APPLICATION NUMBER: 2002-05-03
: PRIOR FILING DATE: 1997-06-18
: PRIOR APPLICATION NUMBER: 60/049911
: PRIOR FILING DATE: 1997-08-26
: PRIOR APPLICATION NUMBER: 60/056974
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059113
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059115
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059117
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059122
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059184
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059263
: PRIOR FILING DATE: 1997-09-18
: PRIOR APPLICATION NUMBER: 60/059352
: PRIOR FILING DATE: 1997-09-19
: PRIOR APPLICATION NUMBER: 60/059588
: PRIOR FILING DATE: 1997-09-19
: PRIOR APPLICATION NUMBER: 60/059836
: PRIOR FILING DATE: 1997-09-24
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17

: PRIOR APPLICATION NUMBER: 60/062285
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/062287
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/062814
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/062816
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063045
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063082
: PRIOR FILING DATE: 1997-10-31
: PRIOR APPLICATION NUMBER: 60/063127
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063327
: PRIOR FILING DATE: 1997-10-27
: PRIOR APPLICATION NUMBER: 60/063329
: PRIOR FILING DATE: 1997-10-27
: PRIOR APPLICATION NUMBER: 60/063550
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063561
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063704
: PRIOR FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: 60/063733
: PRIOR FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: 60/063735
: PRIOR FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: 60/063738
: PRIOR FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: 60/063755
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064248
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/064809
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: 60/065186
: PRIOR FILING DATE: 1997-11-12
: PRIOR APPLICATION NUMBER: 60/065846
: PRIOR FILING DATE: 1997-11-17
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/066453
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/066511
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/066770
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/069212
: PRIOR FILING DATE: 1997-12-11
: PRIOR APPLICATION NUMBER: 60/069278
: PRIOR FILING DATE: 1997-12-11
: PRIOR APPLICATION NUMBER: 60/069334
: PRIOR FILING DATE: 1997-12-11
: PRIOR APPLICATION NUMBER: 60/069694
: PRIOR FILING DATE: 1997-12-16
: PRIOR APPLICATION NUMBER: 60/072320
: PRIOR FILING DATE: 1998-01-23
: PRIOR APPLICATION NUMBER: 60/073612
: PRIOR FILING DATE: 1998-02-04
: PRIOR APPLICATION NUMBER: 60/074086
: PRIOR FILING DATE: 1998-02-09
: PRIOR APPLICATION NUMBER: 60/074092
: PRIOR FILING DATE: 1998-02-09
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-02-27
: PRIOR APPLICATION NUMBER: 60/079728

PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26

PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982

Query Match 67.8%; Score 510.4; DB 14; Length 1266;
Best Local Similarity 79.9%; Pred. No. 2.1e-136;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGACCCCACTGATTCATCTTGGCCACTTCCCTCTGCTCTCAATG 60
DB 10 ATGAGAGGCACACGGAAGACCCTCTGCTTCTCCCTCTGCTCTCAAG 69
QY 61 GTGTGCCCCAGCTGTGCCGACACCTGTACCTGTCTTGACACACCCAGTCCCA 120
DB 70 GTCCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCG 129
QY 121 CAGGGGTACCCCTGTGCTGATGCTGTGCTGTGCTGTAAAGTGTGCACGAGGCTG 180
DB 130 CTGGGAGTACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 189
QY 181 GGGGAGTCTTGGACACCACTGCATGTCTGCGACCCCAAGCCAGGCGCTGTGCTCAAGCT 240
DB 190 GGGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCCAGGCGCTGTGCTCAAGCT 249
QY 241 GGGGAGGCCCCCTGGCGCCATGGGCGCTGTGCTCTCTTGATGAGATGACGCTAGCTGT 300
DB 250 GGGGAGGACCCCGGTGGCGCGGCGCCCTGTGCTCTTGCGAGAGGACGACGACGACGCTGT 309
QY 301 GAGGTGATGCGCCGACGATGCTGATGAGAGACCTTTAAACCAATGACGGTCTGTG 360
DB 310 GAGGTGACGCGCCGCTGTATGCGGAAGGAGAGACCTTCAACCCCACTGACGATCCGC 369
QY 361 TGCCGCTGTGATGACGCTGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 370 TGCCGCTGCGAGGACGCGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429
QY 421 CCCAGTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAGGAAGTGTGCCCGAG 480
DB 430 CCCAGTGGGACTGCCACGCCCCAAGAGAGGTGCCAGGAAGTGTGCCCGAG 489
QY 481 TGGGTATGTGACCAAGGAGTGCACCGCGCATCCAGCGCTCCACGGCGCAAGACACCA 540
DB 490 TGGGTGCGGCGCAAGGAGGCGGAGTGGGACCCAGCCCTTCCAGCCCAAGGACCCAG 549
QY 541 CTTCCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 550 TTTTCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
QY 601 TGGGCCCCCTGCTCAACCACTGTGGCTGGGATAGGACCCAGTGTCCAAACCAAGAC 660
DB 610 TGGGAGCCCTGCTGACCACTGTGGCTGGGATGCGACCCGCGGTGTCCAAACCAAGAC 669
QY 661 CGATTGCGCACTGAGATCCAAAGCGCGCTGTGTGCTGCCAGACCCCTGCTGCGAGCC 720
DB 670 CGCTTGTGCGCACTGAGACCCAGCGCGCTGTGTGCTGCCAGGCGCTGCGCCACCCCTCC 729
QY 721 AGGAGCCACAGCTCATGGAACAGTCTTCTA 752
DB 730 AGGGGTGCGAGTCCACAAACAGTCTTCTA 761

RESULT 7
US-10-146-726-319
Sequence 319, Application US/10146726
Publication No. US20030129690A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen

APPLICANT: Gao,Wei-Qiang
APPLICANT: Gerlitsen,Mary E.
APPLICANT: Goddard,Audrey
APPLICANT: Godowski,Paul J.
APPLICANT: Gurney,Austin L.
APPLICANT: Sherwood,Steven
APPLICANT: Smith,Victoria
APPLICANT: Stewart,Timothy A.
APPLICANT: Tumas,Daniel
APPLICANT: Watanabe,Colin K
APPLICANT: Wood,William
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C308
CURRENT APPLICATION NUMBER: US/10/146,726
CURRENT FILING DATE: 2002-05-15
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 319
LENGTH: 1266
TYPE: DNA
ORGANISM: Homo Sapien
US-10-146-726-319

Query Match 67.8%; Score 510.4; DB 14; Length 1266;
Best Local Similarity 79.9%; Pred. No. 2.1e-136;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCTCAATG 60
Db 10 ATGAGAGGCACACCGAAGACCCACTCTGGCTTCTCTCTCTCTCTCTCTCAAG 69
QY 61 GTGTGTCCCAAGCTGTGCGGAGACCCCTGTACCTGTCTCTTGACACCAACCAGTGCCCA 120
Db 70 GTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCG 129
QY 121 CAGGGGTACCCCTGTGTCTGTATGGCTGTGGCTGTCTTAAGTGTGTGACGAGGCTG 180
Db 130 CTGGAGTACCCCTGTGTCTGTATGGCTGTGGCTGTCTGTATGTGACGAGGCTG 189
QY 181 GGGAGTCTTCGCGACCACTGCATGTCTGCGACCCCAAGGCGCTGTTGTACGCT 240
Db 190 GGGAGCCCTGCGACCACTGCATGTCTGCGACCCCAAGGCGCTGTTGTACGCT 249
QY 241 GGGGAGGCGCTGCGGCGCATGGGCTGTGTCTCTTGGATGAGATGACGATGCTGT 300
Db 250 GGGGAGGACCCGCTGCGGCGCATGGGCTGTGTCTCTTGGATGAGATGACGATGCTGT 309
QY 301 GAGTGAATGCGCGCAGATGATGATGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
Db 310 GAGTGAACGGCCGCTGTATCGGGAAGGGAGAGACCTTCAAGCCCACTGACGATCCGC 369
QY 361 TGGCGCTGTATGACGATGCGCTTCACTGCGCTGCCGCTGTGACATGAGATGTGCGCTG 420
Db 370 TGGCGCTGTATGACGATGCGCTTCACTGCGCTGCCGCTGTGACATGAGATGTGCGCTG 429
QY 421 CCCAGCTGGGACTGCCACGCGCCCAAGAGATACAGGTGCCAGGAAAGTGTGCCCGCAG 480
Db 430 CCCAGCTGGGACTGCCACGCGCCCAAGAGAGGTGAGGTCTTGGGCAAGTGTGCCCTGAG 489
QY 481 TGGGTATGTACCAAGGAGTGAACCGGCGATCCAGCGCTTCAAGGCGCAAGGACACCAA 540
Db 490 TGGGTATGTACCAAGGAGTGAACCGGCGATCCAGCGCTTCAAGGCGCAAGGACACCAA 549
QY 541 CTTCCTGCGCTGTCT 600
Db 550 TTTCTGCGCTGTCT 609
QY 601 TGGGCGCCCTGTCTCAACCACTGTGGGCTGGGCTAGGCAAGGAGTGTCAACAGAGAG 660
Db 610 TGGGAGCCCTGTCTGACCACTGTGGGCTGGGCTAGGCAAGGAGTGTCAACAGAGAG 669

QY 661 CGATTCTGCCAAGTGAATCCACGCGCGCTGTGTCTGCCAGACCCCTGCGTGGAGCC 720
Db 670 CGCTTCTGCCAAGTGAATCCACGCGCGCTGTGTCTGCCAGAGCCCTGCGTGGAGCC 729
QY 721 AGGAGCCACAGCTCATGAGACAGTGTCTCTA 752
Db 730 AGGAGCCACAGCTCATGAGACAGTGTCTCTA 761

RESULT 8
US-10-146-727-319
Sequence 319, Application US/10146727
Publication NO. US20030129691A1
GENERAL INFORMATION:
APPLICANT: Baker,Kevin P.
APPLICANT: Beresini,Maureen
APPLICANT: DeForge,Laura
APPLICANT: Desnoyers,Luc
APPLICANT: Filvaroff,Ellen
APPLICANT: Gao,Wei-Qiang
APPLICANT: Gerlitsen,Mary E.
APPLICANT: Goddard,Audrey
APPLICANT: Godowski,Paul J.
APPLICANT: Gurney,Austin L.
APPLICANT: Sherwood,Steven
APPLICANT: Smith,Victoria
APPLICANT: Stewart,Timothy A.
APPLICANT: Tumas,Daniel
APPLICANT: Watanabe,Colin K
APPLICANT: Wood,William
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C312
CURRENT APPLICATION NUMBER: US/10/146,727
CURRENT FILING DATE: 2002-05-15
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 319
LENGTH: 1266
TYPE: DNA
ORGANISM: Homo Sapien
US-10-146-727-319

Query Match 67.8%; Score 510.4; DB 14; Length 1266;
Best Local Similarity 79.9%; Pred. No. 2.1e-136;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCTCAATG 60
Db 10 ATGAGAGGCACACCGAAGACCCACTCTGGCTTCTCTCTCTCTCTCTCTCAAG 69
QY 61 GTGTGTCCCAAGCTGTGCGGAGACCCCTGTACCTGTCTCTTGACACCAACCAGTGCCCA 120
Db 70 GTGCGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCG 129
QY 121 CAGGGGTACCCCTGTGTCTGTATGGCTGTGGCTGTCTTAAGTGTGTGACAGGAGCTG 180
Db 130 CTGGAGTACCCCTGTGTCTGTATGGCTGTGGCTGTCTGTATGTGACAGGAGCTG 189
QY 181 GGGAGTCTTCGCGACCACTGCATGTCTGCGACCCCAAGGCGCTGTTGTACGCT 240
Db 190 GGGAGCCCTGCGACCACTGCATGTCTGCGACCCCAAGGCGCTGTTGTACGCT 249
QY 241 GGGGAGGCGCTGCGGCGCATGGGCTGTGTCTCTTGGATGAGATGACGATGCTGT 300
Db 250 GGGGAGGACCCGCTGCGGCGCATGGGCTGTGTCTCTTGGATGAGATGACGATGCTGT 309
QY 301 GAGTGAATGCGCGCAGATGATGATGAGAGACCTTTAAACCAATTGACAGGCTCTG 360
Db 310 GAGTGAACGGCCGCTGTATCGGGAAGGGAGAGACCTTCAAGCCCACTGACGATCCGC 369
QY 361 TGGCGCTGTATGACGATGCGCTTCACTGCGCTGCCGCTGTGACAGTGAAGATGTGCGCTG 420

; CURRENT FILING DATE: 2002-05-21
; Prior Application removed - See Palm or File Wrapper.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-152-380-319

Query Match 67.8%; Score 510.4; DB 14; Length 1266;
Best Local Similarity 79.9%; Pred. No. 2.1e-136;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTTCTCTCATG 60
Db 10 ATGAGAGGACACCGAGACCCACCTCTGCGCTCTCTCTCTCTCTCTCTCAAG 69
QY 61 GTGTGTGCCAGCTGTGCGCGAGACACCCTGTACCTGTCTTGACACACCACCCAGTGC 120
Db 70 GTGCGTACCCAGCTGTGCGCGAGACACCATGTACCTGCCCCCTGGCCACCTCCCGATGCCG 129
QY 121 CAGGGGGTACCCCTGTGCTGTGATGGCTGTGCTGTCTTAAAGTGTGTGCAGGAGCTG 180
Db 130 CTGGGAGTACCCCTGTGCTGTGATGGCTGTGCTGTGCTGCGGATGTGTGCAGGCGGCTG 189
QY 181 GGGGAGTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGGCTGTTGTACGCT 240
Db 190 GGGGAGCCCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGGCTGTTGTACGCT 249
QY 241 GGGGAGGCGCTGCGCGCCATGCGGCTGTGCTGTCTTGGATGAGATGAGCGTGTCTGT 300
Db 250 GGGGAGGACCCGCTGCGCGCGCGGCGGCGCTGTGCTGTCTTGGAGAGAGACAGACGCTGT 309
QY 301 GAGGTGAATGCGCGAGTACCTGATGAGAGACCTTTAAACCAATTCAGGGTCTG 360
Db 310 GAGGTGAACGCGCGCTGTATCGGGAAGGAGACCTTCCAGCCCACTGCACATCCGC 369
QY 361 TGCCGCTGTGATGACGGTGGCTTACCTGCTGCTGCTGTGATGAGATGAGGCTG 420
Db 370 TGCCGCTGCGAGAGCGGCGCTTACCTGCTGCTGCTGTGCAAGGATGTGCGCTG 429
QY 421 CCCAGCTGGAGCTGCCACGCCCCCAAGAAATACAGGTGCCAGGAAATGCTGCCCGAG 480
Db 430 CCCAGCTGGAGCTGCCACGCCCCCAAGAGGCTGAGGCTCTGGGCAAGTGTGCTGCCCTGAG 489
QY 481 TGGGTATGTACAGGAGTGCACACCGCGGATCCAGCGCTCCAGCGCGCAAGAGACCAA 540
Db 490 TGGGTGTGCGGCAAGAGGAGGAGTGGGAGCCAGCCCTTCCAGCCCAAGAGACCCAG 549
QY 541 CTTTCTGCCCTGTGATCTCTGCTCTGTGATGCTCTTGTCCAAATTTGAGACAGCC 600
Db 550 TTTTCTGGCCTGTGCTCTTCCCTGCCCCCTGTGCTCTGCCCCAGAAATGAGACAGGCC 609
QY 601 TGGGGCCCTGTCTCAACCACTGTGGCTGGGCAATAGCCACCCGAGTGTCCACCAAGAAC 660
Db 610 TGGGGACCTGTCTCAACCACTGTGGCTGGGCAATAGCCACCCGAGTGTCCACCAAGAAC 669
QY 661 CGATTCTGCCACTGAGATCCACAGCGCGCTGTGTCTGCCAGAGACCTGCTGCCAGCC 720
Db 670 CGCTTCTGCGACTGAGACCCAGCGCGCTGTGTCTGCCAGAGGCGCTGCCACCTGCC 729
QY 721 AGGAGCCACAGCTCATGGAACAGTGTCTTA 752.
Db 730 AGGGGTGCGAGTCCACAAACAGTGTCTTA 761

RESULT 11
US-10-153-934-319

; Sequence 319, Application US/10153934
; Publication No. US20030129695A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C412
; CURRENT APPLICATION NUMBER: US/10/153,934
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-153-934-319

Query Match 67.8%; Score 510.4; DB 14; Length 1266;
Best Local Similarity 79.9%; Pred. No. 2.1e-136;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTTCTCTCATG 60
Db 10 ATGAGAGGACACCGAGACCCACCTCTGCGCTCTCTCTCTCTCTCTCTCAAG 69
QY 61 GTGTGTGCCAGCTGTGCGCGAGACACCCTGTACCTGTCTTGACACACCACCCAGTGC 120
Db 70 GTGCGTACCCAGCTGTGCGCGAGACACCATGTACCTGCCCCCTGGCCACCTCCCGATGCCG 129
QY 121 CAGGGGTACCCCTGTGCTGTGATGGCTGTGCTGTCTTAAAGTGTGTGCAGGAGCTG 180
Db 130 CTGGAGTACCCCTGTGCTGTGATGGCTGTGCTGTGCTGCGGATGTGTGCAGGCGCTG 189
QY 181 GGGGAGTCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGGCTGTTGTACGCT 240
Db 190 GGGGAGCCCTGCGACCACTGCATGTCTGCGACCCAGCCAGGGGCTGTTGTACGCT 249
QY 241 GGGGAGGCGCTGCGCGCCATGCGGCTGTGCTGTCTTGGATGAGATGAGGCTGTCTGT 300
Db 250 GGGGAGGACCCGCTGCGCGCGCGGCGGCGCTGTGCTGTCTTGGAGAGAGACAGACGCTGT 309
QY 301 GAGGTGAATGCGCGAGTACCTGATGAGAGACCTTTAAACCAATTCAGGGTCTG 360
Db 310 GAGGTGAACGCGCGCTGTATCGGGAAGGAGACCTTCCAGCCCACTGCACATCCGC 369
QY 361 TGCCGCTGTGATGACGGTGGCTTACCTGCTGCTGCTGCTGTGAGTGAAGTGTGCGGCTG 420
Db 370 TGCCGCTGCGAGAGCGCGCTTACCTGCTGCTGCTGCTGTGAGCGAGATGTGCGGCTG 429
QY 421 CCCAGCTGGAGCTGCCACGCCCCAAGAGAAATACAGGTGCCAGAGAAATGCTGCCCGAG 480
Db 430 CCCAGCTGGAGCTGCCACGCCCCAAGAGAGGCTGAGGTCTGGGCAAGTGTGCCCTGAG 489
QY 481 TGGGTATGTACAGGAGTGCACACCGCGCATCCAGCGCTCCACGCGCGCAAGAGACCAA 540
Db 490 TGGGTGTGCGGCAAGAGGAGGAGTGGGAGCCAGCCCTTCCAGCCCAAGAGACCCAG 549
QY 541 CTTTCTGCCCTGTGATCTCTGCTCTGTGATGCTCTTGTCCAAATTTGAGACAGCC 600
Db 550 TTTTCTGGCCTGTGCTCTTCCCTGCCCCCTGTGCTCTGCCCCAGAAATGAGACAGGCC 609
QY 601 TGGGGCCCTGTCTCAACCACTGTGGCTGGGCAATAGCCACCCGAGTGTCCAAACAGAAC 660


```
Db      610 TGGGACCCCTGCTCGACCACTGTGGCTGGCATGGCCACCCGGGTGTCACAGAAC 669
OY      661 CGATTCTGCCAACTGAGATCCAAACGCCCTGTGTGTGCCAGACCTGCTGGCAGCC 720
Db      670 CGCTTCTGCCAGCTGAGAGACCAGCGCGCTGTGCGTGTCCAGGCCCTGCCACCTCC 729
OY      721 AGGAGCCACAGCTCATGGAACAGTGTCTTA 752
Db      730 AGGGGTCCAGTCCACAAACAGTGCCTTCTA 761

RESULT 12
US-10-028-072-319
; Sequence 319, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028, 072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
```


APPLICANT: Wood,William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C17
CURRENT APPLICATION NUMBER: US/10/121,049
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 319
LENGTH: 1266
TYPE: DNA
ORGANISM: Homo Sapien
US-10-121-049-319

Query Match 67.8%; Score 510.4; DB 15; Length 1266;
Best Local Similarity 79.9%; Pred. No. 2.1e-136;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

1 ATGAGGGGAGCCCACTGATCCATCTTGGCCACTTGTCTCTCTCTCTCAATG 60
10 ATGAGAGGACACCGAAGACCCACCTCTGCTCTCTCTCTCTCTCTCTCTCAAG 69
61 GTGTGTCCAGCTGTGCGGACACCCCTGTACCTGTCTGTGACACCCAGTGCCTCA 120
70 GTGCGTACCCAGCTGTGCGGACACCACTGTACCTGTCTGTGACACCCAGTGCCTCA 129
121 CAGGGGTAACCCCTGTGCTGTGATGCTGTGCTGTGTAAGTGTGTGACAGGAGCTG 180
130 CTGGAGTACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 189
181 GGGAGTCTCTGCGACCACTGATGCTGTGCGACCCAGCCAGGCGCTGTGTGACGCT 240
190 GGGAGGCTCTGCGACCACTGATGCTGTGCGACCCAGCCAGGCGCTGTGTGACGCT 249
241 GGGGAGGCGCTGTGCGGACCCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 300
250 GGGGAGGCGCTGTGCGGACCCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 309
301 GAGGTGAATGCGCGGACGAGTACCTGTGATGAGAGACCTTTAAACCAATTGACGGTCT 360
310 GAGGTGAATGCGCGGACGAGTACCTGTGATGAGAGACCTTTAAACCAATTGACGGTCT 369
361 TGCCGCTGTGATGAGCGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 420
370 TGCCGCTGTGATGAGCGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 429
421 CCCAGTGGAGCTGCGGACCCAGGAGAGATACAGTGTGAGGAAAGTGTGCGCCGAG 480
430 CCCAGTGGAGCTGCGGACCCAGGAGAGATACAGTGTGAGGAAAGTGTGCGCCGAG 489
481 TGGGTATGTGACAGGAGTGAACCGGCGATCCAGCGCTCCAGCGCGCAAGACCA 540
490 TGGGTATGTGACAGGAGTGAACCGGCGATCCAGCGCTCCAGCGCGCAAGACCA 549
541 CTTTCTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600
550 TTTTCTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 609
601 TGGGGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 660
610 TGGGGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 669
661 CGATTGTGCACTGTGAGATCAACGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 720
670 CGCTTGTGCACTGTGAGATCAACGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 729
721 AGGAGCCAGACTGTGAGATCAACGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 752
730 AGGGGTGCGAGTCCACAAACAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 761

RESULT 14

US-10-123-904-319
Sequence 319, Application US/10123904
Publication No. US20030022328A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Zhang, Zemin
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C54
CURRENT APPLICATION NUMBER: US/10/123,904
CURRENT FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 319
LENGTH: 1266
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-904-319

Query Match 67.8%; Score 510.4; DB 15; Length 1266;
Best Local Similarity 79.9%; Pred. No. 2.1e-136;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

1 ATGAGGGGAGCCCACTGATCCATCTTGGCCACTTGTCTCTCTCTCTCAATG 60
10 ATGAGAGGACACCGAAGACCCACCTCTGCTCTCTCTCTCTCTCTCTCTCAAG 69
61 GTGTGTCCAGCTGTGCGGACACCCCTGTACCTGTCTGTGACACCCAGTGCCTCA 120
70 GTGCGTACCCAGCTGTGCGGACACCACTGTACCTGTCTGTGACACCCAGTGCCTCA 129
121 CAGGGGTAACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 180
130 CTGGAGTACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 189
181 GGGAGTCTCTGCGACCACTGATGCTGTGCGACCCAGCCAGGCGCTGTGTGACGCT 240
190 GGGAGGCTCTGCGACCACTGATGCTGTGCGACCCAGCCAGGCGCTGTGTGACGCT 249
241 GGGGAGGCGCTGTGCGGACCCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 300
250 GGGGAGGCGCTGTGCGGACCCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 309
301 GAGGTGAATGCGCGGACGAGTACCTGTGATGAGAGACCTTTAAACCAATTGACGGTCT 360
310 GAGGTGAATGCGCGGACGAGTACCTGTGATGAGAGACCTTTAAACCAATTGACGGTCT 369
361 TGCCGCTGTGATGAGCGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 420
370 TGCCGCTGTGATGAGCGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 429
421 CCCAGTGGAGCTGCGGACCCAGGAGAGATACAGTGTGAGGAAAGTGTGCGCCGAG 480
430 CCCAGTGGAGCTGCGGACCCAGGAGAGATACAGTGTGAGGAAAGTGTGCGCCGAG 489
481 TGGGTATGTGACAGGAGTGAACCGGCGATCCAGCGCTCCAGCGCGCAAGACCA 540
490 TGGGTATGTGACAGGAGTGAACCGGCGATCCAGCGCTCCAGCGCGCAAGACCA 549

QY	541	CTTCTGCCCCCTGTGCACCTCCCTGCTGATGCTCCCTGTGTCCAAATTGGACACAGCC	600
Db	550	TTTCTGGCCCTGTCTCTTCCCTGCCCCCTGTGTCCTCCCTGCCAAGATGAGCACGGCC	609
QY	601	TGGGGCCCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAGAAC	660
Db	610	TGGGGACCCCTGCTGCACCACCTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAGAAC	669
QY	661	CGATTCTGCCAACTGGAGATCCAACGCCGCCCTGTGTCTGCCACGACCCTGCCCTGGCAGCC	720
Db	670	CGCTTCTGCCGACTGGAGACCCACGCCGCCCTGTGCCCTGTCCAAGGCCCTGCCCAACCCCTCC	729
QY	721	AGGAGCCACACGCTCATGGAACAGTGTCTTCTA	752
Db	730	AGGGGTCCGAGTCCACAAACAGTGCCTTCTA	761

RESULT 15
US-10-140

; Sequence 319, Application US/10140470
; Publication No. US20030022331A1

; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Mauri

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

1. TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

10 TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P33330R1C160

; CURRENT APPLICATION NUMBER: US/10/140,470

CURRENT FILING DATE: 2002-05-06

;
Prior Application removed - See Palm or File Wrapper

NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 319

LENGTH: 1266

; TYPE: DNA

ORGANISM: Homo Sapien

05-10-140-470-319

Query Match	67.8%;	Score 510.4;	DB 15;	Length 1266;
Best Local Similarity	79.9%;	Pred. No. 2.1e-136;		
Matches 601; Conservative	0;	Mismatches 151;	Indels 0;	Gaps 0;

QY 1 ATGAGGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCTCCTCGCTTCCTCAATG 60
 ||||| |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 10 ATGAGAGGCACACCAGAACCACCTCTGGCCCTTCTCCCTCCTCTGCCCTCCCTTCAAG 69

```

OY      61 GTGTGTGCCCAGCTGTGCCCGGACACCCCTGTACCTGTCTTGAGACACCACCCAGTGCCCA 120
        ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      70 GTGCGTACCCAGCTGTGCCCGGACACCATGTACCTGCCCCCTGGCCACCTCCCCGATGCCCG 129

```

OY 121 CAGGGGATACCCCTGTGCTGGATGGCTGTGGCTCTGTAAAGTGTGTGCACCGAGCCTG 180
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 130 CTGGAGTACCCCTGTGCTGGATGGCTGTGGCTGTGCCGGTATGTGCACGGCGGCTG 189

Dy
181 GGGGAGTCTCGGACCACTGCATGTTCGGACCCAGCCAGGGCCTGGTTTGACCT 240

Db
190 GGGGAGCCCTGCGAACCACTCCACTGTGCGAGGCCAGCCAGGGCCTGTTGCCAAGCC 249

[illegible]

Search completed: July 28, 2003, 21:31:59
Job time : 186.622 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 18:02:17 ; Search time 1335.24 Seconds
(without alignments)
9133.322 Million cell updates/sec

Title: US-10-010-408-3
Perfect score: 753
Sequence: 1 ATGAGGGGACGCCACTGAT.....CATGACAGTCTTCTAA 753

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	508.8	67.6	1058	14	BM805088
2	500	66.4	979	14	BQ279131
3	463.6	61.6	886	13	BI822142
4	430.4	57.2	750	13	BM043988
5	430.2	57.1	651	13	BI457141
6	427	56.7	916	13	BI457367

7	421.2	55.9	1006	14	BM921531	BM921531	AGENCOURT
8	417.8	55.5	1073	14	BQ073722	BQ073722	AGENCOURT
9	409	54.3	800	13	BI826781	BI826781	603077268
10	404.2	53.7	928	13	BI161474	BI161474	602864871
11	398.8	53.0	888	13	BI825652	BI825652	603072631
12	387.2	51.4	790	13	BM046275	BM046275	603626068
13	375.2	49.8	1022	14	BQ952960	BQ952960	AGENCOURT
14	365.4	48.5	620	13	BG928868	BQ928868	HNC57-1-D
15	332	44.1	1251	14	BQ961357	BQ961357	AGENCOURT
16	330.8	43.9	1166	13	BM543799	BM543799	AGENCOURT
17	320.8	42.6	489	14	BM751866	BM751866	K-EST0028
18	319.2	42.4	749	9	AL555144	AL555144	AL555144
19	296.8	39.4	933	14	BQ278961	BQ278961	AGENCOURT
20	293.2	38.9	452	10	BE481184	BE481184	166529 BA
21	251	33.3	940	14	BQ937887	BQ937887	AGENCOURT
22	234.4	31.1	657	13	BM488499	BM488499	pgm2n.pk0
23	223.2	29.6	405	12	BG900069	BG900069	H0A51-1-A
24	222	29.5	618	12	BG538695	BG538695	602566932
25	210.4	27.9	537	14	BQ560868	BQ560868	H4067A01-
26	209.6	27.8	380	12	BG900020	BG900020	H0A48-1-G
27	193.8	25.7	424	10	BB849097	BB849097	BB849097
28	193.2	25.7	792	13	BI823598	BI823598	603040962
29	192.6	25.6	401	12	BF849398	BF849398	CM3-EN007
30	171.6	22.8	668	9	AL555143	AL555143	AL555143
31	167.4	22.2	966	12	BF141695	BF141695	601790752
32	162.8	21.6	823	14	BQ067127	BQ067127	AGENCOURT
33	154.4	20.5	948	9	AL542806	AL542806	AL542806
34	153.4	20.4	879	9	AL547439	AL547439	AL547439
35	142	18.9	1076	13	BM554624	BM554624	AGENCOURT
36	135.4	18.0	1002	9	AL556435	AL556435	AL556435
37	132.8	17.6	896	14	BQ721003	BQ721003	AGENCOURT
38	128.2	17.0	1117	12	BF982720	BF982720	602304922
39	127.8	17.0	936	14	BQ880159	BQ880159	AGENCOURT
40	127.4	16.9	887	14	BQ894536	BQ894536	AGENCOURT
41	127	16.9	1057	12	BE898435	BE898435	601681295
42	126.8	16.8	1085	13	BM558306	BM558306	AGENCOURT
43	126.6	16.8	828	9	AL543019	AL543019	AL543019
44	125.2	16.6	657	9	AL711224	AL711224	DKFZp686B
45	125	16.6	505	12	BF037009	BF037009	601456766

ALIGNMENTS

RESULT 1
BM805088
LOCUS
DEFINITION
AGENCOURT_6490429 NIH_MGC_125 Homo sapiens CDNA clone IMAGE:5725937
5', mRNA sequence.
ACCESSION
BM805088
VERSION
BM805088.1 GI:19121911
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 1058)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12717 row: j column: 18
High quality sequence stop: 697.
location/Qualifiers
1. 1058

FEATURES
source


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5725937"
/clone_lib="NIH_MGC_125"
/lab_host="DH10B"
/notes="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcorV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcorV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

```

[illegible]

Query Match	67.6%	Score 508.8;	DB 14;	Length 1058;
Best Local Similarity	79.8%;	Pred. No. 1.2e-114;		
Matches 600; Conservative	0;	Mismatches 152;	Indels 0;	Gaps 0;

OY		1	ATGAGGGGCAGCCCCACTGATTCATCTTCTGGCCACTTCCTTCCTCTGCCCTTCTCAATG	60
Db		9	ATGAGAGGCACACCAGAACCCACTCCTGGCCCTTCTCCCTCTTGCTCCTCTCAAG	68
OY		61	GTCGTGCCAGCTGTGCCCGAGACCCCTGTACCTGTCTTGAGACACCACCATGCCCCA	120
Db		69	GTCGTAACCCAGCTGTGCCCGAGACCATGTACCTGCCCTGGCCACACTCCCAGATGCCG	128
OY		121	CAGGGGGTACCCCTGGTGTGTGAATGGCTGTGGCTGTGTAAAGTGTGTGACGAGGCTG	180
Db		129	CTGGGAGTACCCCTGTGTGTGAATGGCTGTGGCTGTGTGCGGGGTATGTGACGGCGCTG	188
OY		181	GGGGAGTCCCTGCGACCACTGCATGTCTGCGACCCCGAGCGAGGGCCTGTGTTCAGCCT	240
Db		189	GGGGAGCCCTGCGACCACTGCCAGTCTGCGACGCCAGCGAGGGCCTGTGTTCAGCCTC	248
OY		241	GGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTCTTTGGATGAGGATGACGGTAGCTGT	300
Db		249	GGGGCAGGACCCGGTGGACGGGGGGCCCTGTGCTCTTTGGCAGAGGACGACAGCAGCTGT	308
OY		301	GAGGTGAATGGCCGCGAGGTACTGTGATGGAGAGACCTTTAAACCCAATTGCAGGGTCTG	360
Db		309	GAGGTGAACGGCCGCTGTATCGGGAAGGGAGAGACCTTCGAGGCCCACTGCAGCATCCGC	368
OY		361	TGCCGCTGTATGACGAGGTGGCTTCACTGCTGCCGCTGTGACGTGAGGATGTGCGGCTG	420
Db		369	TGCCCTGCGAGGACGGCGGCTTCACTGCTGCCGCTGTGACGAGGATGTGCGGCTG	428
OY		421	CCCAGCTGGGACTGCCACGCCCCAGAAGATAACAGGTGACAGGAAGTGTGCCCCCGAG	480
Db		429	CCCAGCTGGGACTGCCACCCACCCAGAGAGGTCGAGGTGCTGGGCAAGTGTGCCCTGAG	488
OY		481	TGGGTATGTGACCAAGGAGTGAACACGGCGATCCAGCGCTCCACCGGCCCAAGCACCAA	540
Db		489	TGGGTGTGCGGCCAAGGAGGGGGAGCTGGGGACCCAGCCCGTTCAGGCCCAAGGACCCAG	548
OY		541	CTTCTGCCCTGTGCACCTCCTGCTCTGCTGATGCTCTTGTCCAATTGAGCACAGCC	600
Db		549	TTTTCTGGCCTTGTCTCTTCCCTGCCCCCTGTGTCCCTGGCCAGAAATGAGCACGGCC	608
OY		601	TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAC	660
Db		609	TGGGGACCCCTGCTCGAACCACTGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAAGAC	668
OY		661	CGATTTCTGCCAACTGGAGATCCAACGCCGCTGTGTGTGGCCAGACCCCTGGCCAGCC	720
Db		669	CGCTTCTGCCGACTGGAGACCCAGCGCCGCTGTGCTGTCCAGGCGCCCTGCCACCCCTCC	728
OY		721	AGGAGCCACAGCTCATGGAACAGTGTCTTCTA	752
Db		729	AGGGGTCCGAGTCCACAACAACAGTGCCTTCTA	760

RESULT 2
BQ279131
LOCUS
DEFINITION
BQ279131 979 bp mRNA linear EST 07-MAY-2002
AGENCOURT_7046721 NIH_MGC_107 Homo sapiens cDNA clone IMAGE:5805819
5', mRNA sequence.
BQ279131
BQ279131.1 GI:20489339
VERSION
KEYWORDS
EST.
SOURCE
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 979)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

```

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2045 row: k column: 04
High quality sequence stop: 752.
Location/Qualifiers
1. .979

```

```

BASE COUNT      141 a      347 c      313 g      170 t      8 others
ORIGIN
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5805819"
/clone_1b="NIH_MGC_107"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/site_1="EcoRI;
/site_2="XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

Query Match	66.4%;	Score 500;	DB 14;	Length 979;
Best Local Similarity	80.28;	Pred. No. 1.6e-112;		
Matches 587; Conservative	0;	Mismatches 145;	Indels 0;	Gaps 0;

[illegible]

OY		157	TGTAAGTGTGTGCACGGAGCTGGGGGAGTCCTGCGACCACCTGCATGCTTGCGACCC	216
Db		126	TTCCGGGTATGTGCACGGCGGCTGGGGGAGGCCCTGCGAACCACTCCACTCTGCGACGCC	185
OY		217	AGCCAGGGCCTGGTTTGTACAGCTGGGGCAGGCCCTGGCGGCCANTGGGGCTGTGTCTC	276
Db		186	AGCCA - GGCTGTCTGCCAGGCCGGGGCAGGACCCGGGTGACGGGGGCCCTGTGCCTC	244
OY		277	TTGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGACGGTACTTGATGGAGAGACC	336
Db		245	TTGGCAGAGGACGACGACGACTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGAGACC	304
OY		337	TTTAAACCCCAATTGCAGGGTCTGTGCTGCGCTGTGATGACGGTGGCTTACCTGCGTCCG	396
Db		305	TTCCAGCCCCACTGACGATCCGCTGCGCTGCGAGGACGGCGGCTTCACTGCGTCCG	364
OY		397	CTGTGCAGTGAGGATGTGCGGCTGCCAGCTGGGACTGCCACGGCCCCAAGAGAATAACAG	456
Db		365	CTGTGCAGCGAGGATGTGCGGCTGCCAGCTGGGACTGCCCCACCCCAGAGAGGGTCGAG	424
OY		457	GTCACAGGAAGTGTGCCCCGAGTGGGTATGTGACCAAGGAGTGCACACCGCGCATCCAG	516
Db		425	GTCCTGGGCAAGTGTGCCCTGAGTGGGTGTGCGGCCAAGAGGGGAGCTGGGGACCCAG	484
OY		517	CGCTCCACGGCGCAAGGACACAATTTCGCCCTGTCACTCCTGCTGCTGATGCT	576
Db		485	CCCCCTCCAGCCCAAGGACCCCAAGTTTCTGGCCTGTCTCTCTCCCTGCCCCCTGCTGT	544
OY		577	CCTTGTCCAAATTGGAGCACAGCCTGGGGCCCTGCTCAACCACCTGTGGGCTGGGCATA	636
Db		545	CCCTGCCCAAGATGAGACAGCGGCTGGGGACCCCTGCTCGACCACCTGTGGGCTGGGCATG	604
OY		637	GCCACCCGAGTGTCCAACCAAGACCGATTCTGCCAACTGGAGATCCA	683
Db		605	GCCACCCGGGTGTCCAAGCAGAACCGTTCTGCCGACTGGAGACCCA	651

```

/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. | "
BASE COUNT      136 a      305 c      304 g      170 t      1 others
ORIGIN

```

Query Match	56.78%;	Score 427;	DB 13;	Length 916;
Best Local Similarity	79.08%;	Pred. No. 1.4e-94;		
Matches 569; Conservative	0;	Mismatches 146;	Indels 5;	Gaps 5;

QY	31	GCCACTTCCTCCTCGCTTCTCTCAATGGTGTGCCCCAGCTGTGCCGACACCCCTGT	90
Db	1	GCCTTCTCCCTCCTCTGCTCTCTCAAGGTGCGTACCAAGCTGTGCCGACACCATGT	60
QY	91	ACCTGTCTTGACACCAACCCAGTGCCACAGGGGGTACCCCTGTGCTGATGGCTGT	150
Db	61	ACCTGCCCTGG-CACCTCCCCGATGCCCGCTGGAGTACCCCTGTGCTGATGGCTGT	119
QY	151	GGCTGCTGTAAAGTGTGTGCACAGGAGCTGGGGGAGTCTTGCCAGCACCTGCATGTCTGC	210
Db	120	GGCTGCTGCCGGGTATGTGCACGGCGGCTGGGGGAGCCCTGCGACCACTCCACGTCTGC	179
QY	211	GACCCACGACGAGGCGCTGTGTCAGCCTGGGGCAGGCCCTGGCCCATGGGGCTGTG	270
Db	180	GACGCCAGCCA-GGCTGTGCTGTGCCAGCCCGGGGACAGACCCGGTGGACGGGGGCCCTG	238
QY	271	TGCTCTTGGATGAGATGACGGTAGCTGTGAGGTGAATGCGCCGAGGTACCTGGATGA	330
Db	239	TGCTCTTGGCAGAGCAGACAGCAGCTGTGAGGTGAACGGCCGCTGTATCGGAAGG	298
QY	331	GAGACCTTTAAACCAATTGCAGGGTCTGTGCGCGCTGTGATGACGGTGGCTTACCTGC	390
Db	299	GAGACCTTCCAGCCCACTGCAGCATCCGCTGCCGCTGCAGGAGACGGCGCTTACCTGC	358
QY	391	CTGCCGCTGTGCAGTGAAGATGTGCGGCTGCCACGCTGGGACTGCCACGCCCCAAGAGA	450
Db	359	GTGCCGCTGTGCAGCGAGATGTGCGGCTGCCACGCTGGGACTGNCCCCTACCAGAGG	418
QY	451	ATACAGGTGCCAGGAAGTGTGCTGCCCGAGTGGGTATGTGACCAGGAGTGACACCGCG	510
Db	419	GTCGAGTCTTGCGCAAGTGTGCCCTGAGTGGGTGTGCGGCCAAGGAGGGGACTGGG	478
QY	511	ATCCAGCGCTCCAGGGCGCAAGACACCAACTTTCTGCCCCCTGTCACTCTGCTGTGT	570
Db	479	ACCCAGGCGCTTCCAGCCCAAGAACCCCAAGTTTCTGCGCTGTCTTCCCTGCCCCCT	538
QY	571	G-ATGCTCCTTGTCCAATTGGAGCACAGCCTGGGGGCCCTGCTCAACCACCTGTGGCT	629
Db	539	GTGTGTCCCTGCCCAAGATGAGACAGCGGCTGGGGAGCCCTGCTCGA-CACCTGTGGCT	597
QY	630	GGGCATAGCCACCCGAGTGTCCAACAGAACCGATTCTGCCAAGTGGAGATCCAACGCCG	689
Db	598	GGGCATGTACCCCGGTGTCCAACAAGAACCGTTCTGCGCACTGGAGACCCAGCGCG	657
QY	690	CC-TGTGTCTGCCACGACCTGCTGGCAGCCAGAGAGCCAGCTCATGGAACAGTGTCT	748
Db	658	CCTTGTGCTGTCCAGGCGCTGACACAGCCTCCAGGGGTGCGAGTCCACACAACAGTGTCT	717
RESULT 7			
LOCUS	BM921531	1006 bp	mRNA linear EST 12-MAR-2002
DEFINITION	AGENCOURT_6708025 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753009		
ACCESSION	BM921531		
VERSION	BM921531.1	GI:19371910	
KEYWORDS	EST.		

SOURCE ORGANISM	REFERENCE TITLE	AUTHORS	JOURNAL	COMMENT
human.				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
1 (bases 1 to 1006)				
NIH-MGC	http://mgc.ncl.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished (1999)				
Contact: Robert Strausberg, Ph.D.				
Email: cgabbs-remail.nih.gov				
Tissue Procurement: Life Technologies, Inc.				
CDNA Library Preparation: Life Technologies, Inc.				
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
DNA Sequencing by: Agencourt Bioscience Corporation				
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:				
http://lml.nih.gov				
plate: L1A12788				
row: b				
column: 18				
High quality sequence stop: 593.				
location/Qualifiers				
1. 1006				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/clone="IMAGE:5753009"				
/clone_1db="NIH_MGC_115"				
/lab_host="DH10B"				
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."				
BASE COUNT	148 a	368 c	317 g	169 t
ORIGIN				4 others
Query Match	55.9%	Score 421.2;	DB 14;	Length 1006;
Best Local Similarity	77.5%;	Pred. No. 3.9e-93;		
Matches	585;	Conservative	0;	Mismatches 161; Indels 9; Gaps 6;
OY	8	GCAGCCCACTGATCCATCTTGCGCACTTCCTCCCTCTCTCATGCTGTG	67	
Db	1	GCACACCGAAGACCCACCTCTGCGCTTCTCCCTCTCTCAAGTGCCTA	60	
OY	68	CCCAAGCTGTGCGGACACCTGTACTGTCTTGACACACCCCAAGTGCCTG	127	
Db	61	CCCAAGCTGTGCGGACACCTGTACTGTCTTGACACCTGCGGATGCGG	120	
OY	128	TACCCCTGTGCTGGATGCTGTGCTGTGTAAGTGTGTCACGAGGCTGGG	187	
Db	121	TACCCCTGTGCTGGATGCTGTGCTGTGCTGCGGATGTCACGAGGCTGGG	180	
OY	188	CCTGCGACCACTGCATGCTGCGACCCAGCCAGGCGCGGTTGTACGCTGGG	247	
Db	181	CCTGCGACCACTGCATGCTGCGACCCAGCCAGGCGCGGTTGTACGCTGGG	240	
OY	248	GCCCTGGCGGCATGGGCTGTGTCTCTTGATGAGATGACGGTAGCTGAG	307	
Db	241	GACCCGGTGGACGGGGGCGCTGTGCTCTTGAGAGAGGAGACAGCAGCTG	300	
OY	308	ATGCGCCGAGGTACCTGCATGAGAGAGACCTTAAACCCAAATTGACAGGCT	367	
Db	301	ACGGCCGCTGTATCGGAGAGGGGAGACCTTCCAGCCCACTGCAGCATCCG	360	
OY	368	GTGATGACGGTGTGCTTCACTGCTGCGGCTGTGACGAGATGCGGCTGCC	427	
Db	361	GCGAGGACGGCGCTTCACTGCTGCGGCTGTGACGAGATGCGGCTGCC	420	
OY	428	GGGACTGCCACGCCCCAAGAGATACAGGTGCAGGAAAGTGTGCCCCGAGTGG	487	

Db 421 GGGACTGCCCCACCCCCAGAGG6GTGAGGTCTCTGGGCAAGTCTGCCCTGAGTGGGTGT 480

QY 488 GTGACCAGGGAGTGAACACCGCGATCCAGCGCTCCACGGCGCAAGACAACCAACTTTCTG 547
| | | | | | | | | | | | | | | | | | | | | |
Db 481 GCGGCCAAGAGGAGGGGACTGGGGACCCAGCCCCTTCACGCCAAGAACCCCAAGTTTTCTG 540

QY 548 CCCTGTGCACCTCCTGCCTCTGCTGATGCTCCTTGTCCAATTGGAGCAC-AGCCTGGGGC 606
| | | | | | | | | | | | | | | | | | | | | |
Db 541 GCCTGTCTCTTCCCTGCCCTGTGTCTCCCTGCCCCAGAATGAGCACAGNGCCTGGGGA 600

QY 607 CCTGCTCAACCACC--TGTGGGCTGGGCATAG-CCACCCGAGTGTCCAA--CCAGAACC 661
| | | | | | | | | | | | | | | | | | | | | |
Db 601 CCTGTCTGACCAACCCNTGTGGCTGNGCATGGGCCACCCGGGTGTCCAACCAAGAACCN 660

QY 662 GATTCCTGCCAACTGGAGATCCAAAGS-CGCGCTGTGTGTGCCAGACCCTGCGGACGC 720
| | | | | | | | | | | | | | | | | | | | | |
Db 661 GCTTCTGCCGACTGGAGACCAGCGCCCGCTGTGCTGTGCCAAGGCGCTGCCCCCTCC 720

QY 721 A--GGAGCCACAGCTCATGGAACAGTGTCTTCTAA 753
| | | | | | | | | | | | | | | | | | | | | |
Db 721 AGGGGGGTGCGAGTCCCCCAAACAGTGCCTTCTAA 755

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	BASE COUNT	ORIGIN
8	BO073722	BO073722 1073 bp mRNA linear EST 02-APR-2002	AGENCOURT_7046577	NIH_MGC_101	Homo sapiens	CDNA clone	IMAGE:5806602								
		5', mRNA sequence.	BO073722												
		EST.	BO073722.1	GI:19902768											
		human.													
		Homo sapiens													
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;													
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.													
		1 (bases 1 to 1073)													
		NIH-MGC http://mgc.ncl.nih.gov/ .													
		National Institutes of Health, Mammalian Gene Collection (MGC)													
		Unpublished (1999)													
		Contact: Robert Strausberg, Ph.D.													
		Email: cgabs-r@mail.nih.gov													
		Tissue Procurement: ATCC													
		CDNA Library Preparation: Rubin Laboratory													
		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)													
		DNA Sequencing by: Agencourt Bioscience Corporation													
		Clone distribution: MGC clone distribution information can be													
		found through the I.M.A.G.E. Consortium/LLNL at:													
		http://image.llnl.gov													
		plate: LLCM2047 row: k column: 19													
		High quality sequence stop: 517.													
		Location/Qualifiers													
		1..1073													
		/organism="Homo sapiens"													
		/db_xref="taxon:9606"													
		/clone="IMAGE:5806602"													
		/clone_lib="NIH_MGC_101"													
		/tissue_type="epidermoid carcinoma, cell line"													
		/lab_host="DH10B (phage-resistant)"													
		/note="Organ: Lung; Vector: pOTB7; Site_1: EcoRI; Site_2:													
		XhoI; cDNA made by oligo-dT priming. Directionally cloned													
		into EcoRI/XhoI sites using the following 5' adaptor:													
		GGCAGCAG(G). Library constructed by Ling Hong in the													
		laboratory of Gerald M. Rubin (University of California,													
		Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and													
		Superscript II RT (Life Technologies). Note: this is a													
		NIH_MGC Library."													
		166 a 384 c 347 g 176 t					</								

Db	668	CCAGCGCGCTGTGCTGTCCAGGCGCCCTGGCCACCTCCAGGGGTGCGAGTCCAGAAA	727
RESULT 10			
LOCUS	BI161474	928 bp	mRNA linear EST 05-JUL-2001
DEFINITION	602864871F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5019065 5',		
ACCESSION	BI161474		
VERSION	BI161474.1	GI:14621475	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 928)		
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaphs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Ling Hong/Rubln Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	plate: L1CM1833 row: m column: 18		
	High quality sequence start: 28		
	High quality sequence stop: 756.		
FEATURES	Location/Qualifiers		
source	1..928		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5019065"		
	/clone_1lb="NIH_MGC_42"		
	/tissue_type="epithelioid carcinoma cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;		
	Site_2: EcoRI; cDNA made by oligo-dT priming.		
	Directionally cloned into EcoRI/XhoI sites using the		
	following 5' adaptor: GGCACGAG(G). Size-selected >500bp		
	for average insert size 1.8kb. Library constructed by Ling		
	Hong in the laboratory of Gerald M. Rubin (University of		
	California, Berkeley) using ZAP-cDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies).		
	Note: this is a NIH_MGC Library. "		
BASE COUNT	161 a 301 c 309 g 157 t		
ORIGIN			
Query Match	53 7%; Score 404.2; DB 13; Length 928;		
Best Local Similarity	76.2%; Pred. No. 5.8e-89;		
Matches	551; Conservative 0; Mismatches 163; Indels 9; Gaps 4;		
QY	21 CCATCTTCTGGCCACTTCCTTCCTCTGCGCTTCTCTCAATGCTGTGCCCCAGCTGTGCCG	80	
Db	82 CCACCTCTGCGCTTCTTCCTCTCTGCGCTCTCTCAAAAGGTGCGTACCCAGCTGTGCC	141	
QY	81 GACACCCCTGTACCTGTCTTGGACACACACCCAGTGGCCACAGGGGGTACCCTGTGCT	140	
Db	142 GACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCGCTGGAGTACCCTGTGCT	201	
QY	141 GGATGGCTGTGGCTGCTGTAAGTGTGTGCACAGGAGCGTGGGGAGTCTCGCACCACT	200	
Db	202 GGATGGCTGTGGCTGCTGCGGATGTGTGCACAGGCGGCTGGGGAGGCCCTGCGACCACT	261	
QY	201 GCATGCTGTGGACCCAGCCAGGGCGCTGTTGTGACGCTGGGGCAGGCCCTGGCGCCA	260	
Db	262 CCACGTCTGCGACGCCAGCCAGGGCGCTGTGTGCCAGCCCGGGCAGGACCCGTTGACG	321	
QY	261 TGGGGCTGTGTCTCTTGGATGAGGATGACGGGTAGCTGTGAGTGAATGCGCGAGTA	320	

Db	322	GGGGGCCCTGTGCCTCTCTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGGCCCTGTA	381
QY	321	CCTGGATGAGAGACCTTTAAACCAATTGACGGTCTGTGCCGCTGTGATGACGGTGG	380
Db	382	TCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGATCCGCTGCCGTGGAGGACGGCGG	441
QY	381	CTTCACCTGCGCTGCCGCTGTGCAAGTGAAGATGTGCGGCTGCCACAGCTGGGACTGCCACG	440
Db	442	CTTCACCTGCGCTGCCGCTGTGCAAGCAGGAGATGTGCGGCTGCCACAGCTGGGACTGAACCCA	501
QY	441	CCCCAAGAGATACAGGTGCCAGGAAGTGTGCCCCGAGTGGGTATGTGACAGGGAGT	500
Db	502	ACCCAGAGAGGTGAGGTCTGCGCAAGTGTGCTGCTGAGTGGGTGTGCGGCCAAGGAGG	561
QY	501	GACACCGGCGATCCAGCGCTCCACAGCGCGCAAGAGACCAACTTCTGCCCCCTGTCACT--	558
Db	562	GGGACTGGGAGACCCAGACCTTACAGAGCCCAAGACCCAGATTCTGGCCTGTCTCTTC	621
QY	559	CCTGCCTCTGCTGATGCTCTCTTGTCCAAATTTGAGCACAGCCTGGGG--CCCTGCTCAAC	617
Db	622	CCATGCCCCACTGTGTCTCCCTGACACAGATGAGACACGGCTTGGGGAACCCCTGCTCGAC	681
QY	618	CACCTGTGGG--CTGGGCATAGCCACCCGAGTGTCCAAACAGAACCGATTCTCCAACTG	675
Db	682	CACATGTGGGGCTGGGCAATTGGACACACCGGGGTGTCCCAACAGAACCGATTCTGCCGACTG	741
QY	676	GAGATCCAA---CGCCGCTGTGTCTGCCACAGCCTGCTGGCAGCCAGAGCCACAG	731
Db	742	GAGACCCAGCGCCGCTAGTGGCCTGTCTCAGGCCCTGCCAACCTCCAGGGGTGCGAA	801
QY	732	CTC	734
Db	802	GTC	804
RESULT 11			
BI825652		888 bp	mRNA
LOCUS	603072631F1	NIH_MGC_119	Homo sapiens cDNA clone IMAGE:5164614 5',
DEFINITION		mRNA sequence.	
ACCESSION	BI825652		
VERSION	BI825652.1	GI:15937202	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 888)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	plate: LLAM11408 row: n column: 07		
	High quality sequence stop: 877.		
FEATURES			
Source	1. 888		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5164614"		
	/clone_lib="NIH_MGC_119"		
	/tissue_type="medulla"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;		
	Site_2: EcoRV (destroyed); RNA source normal medulla from		
	anonymous male age 27. Library is oligo-dT primed and		

Db 422 TGCCCCCTGGTGTCCCTGCCCCAGATGAGCAGCGCCTGGGGACCTGCTCGACCACCT 481

QY 623 GTGGGCTGGGCATAGCCACCCGAGTGTCCCAACCAAGACCGATTCGCCAAGTGAGATCC 682

Db 482 GTGGGCTGGGCATAGCCACCCGAGTGTCCCAACCAAGACCGCTTCGCCAAGTGAGATCC 541

QY 683 AACGCCCTGTGTCTGCCAGACCCCTGGCCAGCCAGCCACAGCTCATGGAACA 742

Db 542 AGCGCCGCTGTGCTGTCCAGGCGCTGCCACCCTCCAGGGGTGCGAGTCACAAACA 601

QY 743 GTGCTTCTA 752

Db 602 GTGCTTCTA 611

RESULT 13

BQ952960 1022 bp mRNA linear EST 21-AUG-2002

LOCUS AGENCOURT_8931684 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6484123

DEFINITION 5', mRNA sequence.

ACCESSION BQ952960

VERSION BQ952960

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1022)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTP

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Plate: LCM2669 row: m column: 20

High quality sequence stop: 482.

FEATURES

source location/Qualifiers

1..1022

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6484123"

/clone_1lb="NIH_MGC_40"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 129 a 305 c 457 g 131 t

ORIGIN

Query Match 49.8%; Score 375.2; DB 14; Length 1022;

Best Local Similarity 78.1%; Pred. No. 8.1e-82;

Matches 467; Conservative 0; Mismatches 123; Indels 8; Gaps 1;

QY 152 GCTGCTTAAGTGTGTGCACGAGCGTGGGGAGTCTGTGCACACCTGCATGTGCG 211

Db 1 GCTGCTGCCGGGTATGTGCACGCGGCTGGGGAGCCGTGCCACCACTCCACGCTGCG 60

QY 212 ACCCCAGCAGGGCCTGTTGTCAAGCTGGGGCAGGGCCTGGGCCCATGGGCTGTGT 271

Db 61 ACGCCAGCAGGGCCTGTTGTGCACGCGGCGCAGGACCCGCGTGGCCGGGGCCCTGT 120

QY 272 GTCTCT-----TGATGAGATGACGGTAGCTGTGAGTGAATGGCCGAGTACCT 323

Db 121 GCCTCTTAAGCAGTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGGCCCTGTATCG 180

QY 324 GGATGAGAGACCTTTAAACCCCAATTGCCAGGGTCTGTGCCCTGTGATGACGGTGGCTT 383

Db 181 GGAAGGGAGACCTTCCAGCCCCCAGCTGCAGCATCCGCTGCCCTGGCAGACGGCGCTT 240

QY 384 CACCTGCTCCCGCTGTGCAGTGAAGATGCGGCTGCCAGCTGGGACTGCCACGCC 443

Db 241 CACCTGCTCCCGCTGTGCAGCGAGATGCGGCTGCCAGCTGGGACTGCCACGCC 300

QY 444 CAAGAGATACAGGTGCCAGGAAGTCTGCCCGAGTGGGTATGTGACAGGAGTGAAC 503

Db 301 CAGAGGGTGCAGGTCTCTGGGCAAGTGTGCCCTGAGTGGGTGCGGCCAAGGAGGGG 360

QY 504 ACCGGCATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCCCTGTCACTCCTGC 563

Db 361 ACTGGGACCCACGCCCCCTTCCAGCCCAAGGACCCAGTTTCTGGCCTGTCTCTCCCT 420

QY 564 CTCTGCTGATGCTCCTTGTCCAAATTGGAAGCAGACAGCCTGGGGCCCCCTGCTCAACCACTG 623

Db 421 GCCCCTGTGTCTCCCTGCCCCAGATGGAAGCAGCGCCTGGGAGCCCTGCTCGACCACCTG 480

QY 624 TGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGACCGATTCGCCAAGTGAGATCCA 683

Db 481 TGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGACCGCTTCTGCGACTGAGACCCA 540

QY 684 ACGCCGCTGTGTCTGCCAGACCTGCTGGCAGCCAGGACCAAGCTCATGGAAC 741

Db 541 GCGCGCTGTGTCTGTCCAGGCGCTGCCACCCCTCCAGGGGTGCGAGGCCCCCAAC 598

RESULT 14

BG928868 620 bp mRNA linear EST 06-NOV-2001

LOCUS HNC57-1-D9.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA

DEFINITION sequence.

ACCESSION BG928868

VERSION BG928868

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 620)

AUTHORS Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J., Sathe, G., Mui, P., Agarwal, P., Badger, A.M., Lee, J.C., Gowen, M. and Lark, M.W.

TITLE Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries

JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)

MEDLINE 21482651

COMMENT Contact: Sanjay Kumar

UW2109

GlaxoSmithKline

709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA

Tel: 610-270-7245

Fax: 610-270-5598

Email: sanjay_kumar-l@sk.com

Seq primer: T7.

FEATURES

source location/Qualifiers

1..620

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_1lb="HNC (Human Normal Cartilage)"

/tissue_type="cartilage"

/lab_host="E.coli DH10 B"

/note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI; Directional"

BASE COUNT 97 a 218 c 207 g 98 t

ORIGIN

Query Match 48.5%; Score 365.4; DB 13; Length 620;
Best Local Similarity 80.2%; Pred. No. 1.8e-79;
Matches 429; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY	3	GAGGGGACAGCCACTGATCCATCTTCTGGCCACTTCCTTCTGCTTCTCAATGCT	62
Db	86	GAGAGGCACACACCGAAGACCACCTCTGCGCTTCTCTCTGCTCTCAAGGT	145
QY	63	GTTGCCCCAGCTGTGCGGACACCTGTACCTGTCTTGACACACCCAGTGGCCCA	122
Db	146	GCGTACCCAGCTGTGCGGACACATGTACCTGCCCTGGCCACCTCCCGATGCCGT	205
QY	123	GGGGGTACCCCTGTGTGATGCTGTGCTGTGTAAGTGTGTGACGAGGCTGG	182
Db	206	GGGAGTACCCCTGTGTGATGCTGTGCTGTGCTGCGGTATGTGACGCGGCTGG	265
QY	183	GGAGTCTGCGACCACTGATGTCTGCGACCCCAAGGCGCTGTTGTACGCTGG	242
Db	266	GGAGCCCTGCGACCACTGCTGTGCGACGCGGCGCTGTGCTGCGACGCGCG	325
QY	243	GGCAGGCCCTGGCGGCAATGGGCTGTGTCTCTTGATGAGATGACGGTAGCTGA	302
Db	326	GGCAGGACCCGGTGGCGGGGCGCTGTGCTCTTGCGAAGGACGACAGCAGCTGTA	385
QY	303	GGTGAATGGCCGCGAGGTACCTGATGGAGAGACCTTTAAACCAATTGCAAGGCTCTG	362
Db	386	GGTGAACGGCGCGCTGTATCGGGAAGGGAGACCTTCAGCCCACTGACATCCGCTG	445
QY	363	CCGCTGTGATGACGGTGTCTTCACTGCTGCCGTGTGCGAGTGAAGTGTGCGCTGCC	422
Db	446	CCGCTGTGAGGACGGCGGCTTCACTGCTGCCGTGTGCGAGGATGTGCGGCTGCC	505
QY	423	CAGCTGGGACTGCCCCACGCCCCAGAGAATACAGGTGCGCAAGAAAGTGTGCCCCGAGTG	482
Db	506	CAGCTGGGACTGCCCCACGCCCCAGAGAGGTGCGAGTCTGCGCAAGTGTGCCCCGAGTG	565
QY	483	GGTATGTGACCAAGGAGTGAACCGGCGATCCAGCGCTCCACGGCGCAAGGACAC	537
Db	566	GGTGTGCGGGCAAGGAGGGGAGTGGGAGCCAGCCCTTCCAGCCCAAGGACCC	620

RESULT 15
BQ961357 1251 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8929398 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6484532
DEFINITION 5', mRNA sequence.

ACCESSION BQ961357
VERSION BQ961357.1 GI:22376835
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1251)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM2670 row: n column: 21
High quality sequence stop: 465.

FEATURES
Source
1.1251
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:6484532"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 178 a 393 c 546 g 131 t 3 others
ORIGIN

QY	1	ATGAGGGGACAGCCACTGATCCATCTTCTGGCCACTTCCTTCTGCTTCTCAATG	60
Db	14	ATGAGAGGCACACCGAAGACCACCTCTGCGCTTCTCTCTGCTCTCAAGGT	73
QY	61	GTTGTTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGACACACCCAGTGGCCA	120
Db	74	GTTGTTGCCAGCTGTGCGGACACCATGTACCTGCGCTGCGCACCTCCCGATGCCG	133
QY	121	CAGGGGGTACCCCTGTGTGATGCTGTGCTGTGCTGTGTAAGTGTGTGACGAGCTG	180
Db	134	CTGCGAGTACCCCTGTGTGATGCTGTGCTGTGCTGTGCTGCGGATGTGACGCGGCTG	193
QY	181	GGGAGTCTCTGACCACTGATGTCTGCGACCCCAAGGCGCTGTTGTGACCT	240
Db	194	GGGAGCCCTGCGACCACTGATGTCTGCGACGCGGAGGCGCTGTGCGACGCC	253
QY	241	GGGCGAGCCCTGCGGCGCATGGGCGCTGTGTCTCTTGATGAGATGACGCTGT	300
Db	254	GGGCGAGGACCCGGTGGACGGGGGCGCTGTGCTCTTGCGAGAGGACGACGAGCTGT	313
QY	301	GAGTGAATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATTGCAGGGTCTG	360
Db	314	GAGTGAACGGCGCGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGACGATCCGC	373
QY	361	TGCGGCTGTGATGACGGTGTGCTTCACTGCTGCCGCTGTGCAAGTATGCGGCTG	420
Db	374	TGCGGCTGTGAGGACGGCGGCTTCACTGCTGCCGCTGTGCAAGGATGTGCGGCTG	433
QY	421	CCGAGCTGGGACTGCCACGCCCCAAGAGAATACAGGTGCCAGGAAAGTGTGCCCG	478
Db	434	CCGAGCTGGGACTGCCACGCCCCAAGGAGGCTCNAAGGGCGGNGCAAGGCTGCCCTG	491

Search completed: July 29, 2003, 02:51:36
Job time : 1337.44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:19:16 ; Search time 4746.55 Seconds
(without alignments)
10472.371 Million cell updates/sec

Title: US-10-010-408-1
Perfect score: 1708
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GCCCTAGATAAACACCCCAA 1708

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query				DB	ID	Description
No.	Score	Match	Length				
1	611	35.8	1741	10	AF259981	AF259981 Rattus no	
2	382	22.4	137964	2	AC126895	AC126895 Rattus no	
3	369	21.6	226303	2	AC095418	AC095418 Rattus no	
4	166	9.7	137964	2	AC126895	AC126895 Rattus no	
5	90	5.3	1734	6	AR210324	AR210324 Sequence	
6	90	5.3	1734	6	AR210325	AR210325 Sequence	
7	90	5.3	1734	10	AF100778	AF100778 Mus muscu	
8	90	5.3	61072	10	AL731698	AL731698 Mouse DNA	
9	90	5.3	216757	2	AL669906	AL669906 Mus muscu	
10	80	4.7	1739	10	AF126063	AF126063 Mus muscu	
11	32	1.9	738	6	AR210337	AR210337 Sequence	
12	32	1.9	841	6	AR210338	AR210338 Sequence	
13	32	1.9	1266	6	AX076919	AX076919 Sequence	
14	32	1.9	1266	6	AX464186	AX464186 Sequence	
15	32	1.9	1283	9	AF083500	AF083500 Homo sapi	
16	32	1.9	1293	6	AR210322	AR210322 Sequence	
17	32	1.9	1293	6	AR210323	AR210323 Sequence	
18	32	1.9	1309	9	AF074604	AF074604 Homo sapi	
19	32	1.9	1427	9	AF100780	AF100780 Homo sapi	
20	32	1.9	1450	9	BC017782	BC017782 Homo sapi	
21	32	1.9	107260	9	AL139352	AL139352 Human DNA	
22	27	1.6	51	6	AR210371	AR210371 Sequence	
23	27	1.6	51	6	AX076923	AX076923 Sequence	
24	23	1.3	145540	2	AC015962	AC015962 Homo sapi	
25	23	1.3	145897	2	AC108665	AC108665 Rattus no	
26	23	1.3	176107	2	AP002393	AP002393 Homo sapi	
27	23	1.3	176612	2	AC090246	AC090246 Homo sapi	
28	23	1.3	185568	9	AC105227	AC105227 Homo sapi	
29	22	1.3	11621	5	AB057676	AB057676 Oryzias l	
30	22	1.3	99395	2	AC010446	AC010446 Homo sapi	
31	22	1.3	142142	2	AC091173	AC091173 Homo sapi	
32	22	1.3	149483	2	AC110904	AC110904 Mus muscu	
33	22	1.3	180464	2	AC116351	AC116351 Homo sapi	
34	22	1.3	186676	9	AC067881	AC067881 Homo sapi	
35	22	1.3	200050	1	AL646068	AL646068 Ralstonia	
36	21	1.2	66908	2	AC119914	AC119914 Mus muscu	
37	21	1.2	150944	2	AC112856	AC112856 Rattus no	
38	21	1.2	154881	2	AC122102	AC122102 Rattus no	
39	21	1.2	167691	2	AC103349	AC103349 Mus muscu	
40	21	1.2	173702	2	AC114696	AC114696 Rattus no	
41	21	1.2	179236	2	AC107112	AC107112 Rattus no	
42	21	1.2	181343	10	AL671882	AL671882 Mouse DNA	
43	21	1.2	204937	2	AL831741	AL831741 Mus muscu	
44	21	1.2	215105	2	AC073717	AC073717 Mus muscu	
45	21	1.2	217917	10	AL672068	AL672068 Mouse DNA	

ALIGNMENTS

RESULT 1
AF259981
LOCUS AF259981 1741 bp mRNA ROD 09-MAY-2000
DEFINITION Rattus norvegicus CCN family protein COP-1 (COP-1) mRNA, complete
cds.
ACCESSION AF259981
VERSION AF259981.1 GI:7739780
KEYWORDS Rattus norvegicus.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1741)
Zhang,R., Averboukh,L., Zhu,W., Zhang,H., Jo,H., Dempsey,P.J.,

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL
Coffey,R.J., Pardee,A.B. and Liang,P.							
Identification of rCop-1, a new member of the CCN protein family,							
as a negative regulator for cell transformation							
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)							
98414629							
9742130							
2 (bases 1 to 1741)							
Liang,P.							
Direct Submission							
Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer							
Center, 649 MRB II, Nashville, TN 37232, USA							
Location/Qualifiers							
1..1741							
/organism="Rattus norvegicus"							
/db_xref="taxon:10116"							
1..1741							
/gene="Cop-1"							
262..1014							
/gene="Cop-1"							
/note="secreted protein"							
/codon_start=1							
/product="CCN family protein COP-1"							
/protein_id="AAF69011.1"							
/db_xref="GI:7739781"							
/translation="MRGSPILRLATSPFLGLSMVCAQLCRTPTCPMPBPQPOGVPLVLDCGCCCKVCARRLTESCEHLHVCPSGLVCOPGAGPGHGAVCLDEDDCEVNGRRYLDGETFEKPNCRVLCRCDDGGFTCLBCSEDTLPMSDCPRPKRIYBGKCPPEWCDQGYTPAIQRSAQGHLSALVTPASADAPWPWNSTAWGPCSTTCGLIATRVSNNQNRFQLEIQIRLCLPRLCLARSHSWSNAF"							
BASE COUNT	386	a	491	c	480	g	384
ORIGIN							t
Query Match	35.8%	Score	611;	DB	10;	Length	1741;
Best Local Similarity	99.4%;	Pred.	No.	0;			
Matches	861;	Conservative	0;	Mismatches	5;	Indels	0;
Gaps	0;						
OY	665	GCTGCCACGCTGGGACTGCCACCGCCCCAAGAGATACAGGTGCCAGGAAGTCTGCC	724				
Db	678	GCTGCCACGCTGGGACTGCCACCGCCCCAAGAGATACAGGTGCCAGGAAGTCTGCC	737				
OY	725	CGAGTGGGTATGTACCAGGAGTGACACCGGCATCCAGCGCTCCACGGCGCAGGACA	784				
Db	738	CGAGTGGGTATGTACCAGGAGTGACACCGGCATCCAGCGCTCCACGGCGCAGGACA	797				
OY	785	CCAACCTTTCTGCCCCTGTCACTCTGCTCTGCTGATGCTCTTGTCCAAATTGACAC	844				
Db	798	CCAACCTTTCTGCCCCTGTCACTCTGCTCTGCTGATGCTCTTGTCCAAATTGACAC	857				
OY	845	AGCCTGGGGCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGTGTCAAACA	904				
Db	858	AGCCTGGGGCCCTGCTCAACCACCTGTGGGCTGGGCATAGCCACCCGAGTGTCAAACA	917				
OY	905	GAACCGATTCTGCCAATGAGATCCAACGCGCGCTGTCTCTGCCAGACCCTGCCCTGCC	964				
Db	918	GAACCGATTCTGCCAATGAGATCCAACGCGCGCTGTCTCTGCCAGACCCTGCCCTGCC	977				
OY	965	AGCCAGGAGCCACAGCTCATGGAACAGTGTCTTTTAAGGCCAACTGGGGATGCGGATACA	1024				
Db	978	AGCCAGGAGCCACAGCTCATGGAACAGTGTCTTTTAAGGCCAACTGGGGATGCGGATACA	1037				
OY	1025	GGGCTGCCATCTCAGCAAAATGACCCTAGACCAAGGCCCTGGACTGCTGTAGATGCTC	1084				
Db	1038	GGGCTGCCATCTCAGCAAAATGACCCTAGACCAAGGCCCTGGACTGCTGTAGATGCTC	1097				
OY	1085	TTCTCCATGCTCTTGCGCTGACAGTTAACTGTCTGCTTGGATTCACTGTGTAGAGCACTG	1144				
Db	1098	TTCTCCATGCTCTTGCGCTGACAGTTAACTGTCTGCTTGGATTCACTGTGTAGAGCACTG	1157				
OY	1145	AGCGATCCCTGCTCTGTCTGAGGTAGGCGGAGCAGGTGACCAAGCTCCAGTTCTCTGGTTTC	1204				
Db	1158	AGCGATCCCTGCTCTGTCTGAGGTAGGCGGAGCAGGTGACCAAGCTCCAGTTCTCTGGTTTC	1217				
OY	1205	AGCCTGGAAATCTGGGTTCTCCTGGCTCAATTCCTCAAAAACATCCCTGTACAAAAAGACA	1264				

Db	1218	AGCGTGAATCTTGGGGTTCCTCCTGCTCATCTCCCAAAACATCCCCTGTACAAGAAGACA	1277	
QY	1265	ACCAAAAAGACCCTTTAACCCTAGGCTATACTGGGCAAACCTGGCCACCCTGCTGGGATA	1324	
Db	1278	ACCAAAAAGACCCTTTAAACCTAGGCTATACTGGGCAAACCTGGCCACCCTGCTGGGATA	1337	
QY	1325	AGGTCATGTTAGGACCAGACAGAGATTGCCCTGAAAACTTCCAATTCCTCTTGGACTT	1384	
Db	1338	AGGTCATGTTAGGACCAGACAGAGATTGCCCTGAAAACTTCCAATTCCTCTTGGACTT	1397	
QY	1385	CTGTATGCTTTGTCGCCCAAGATGATGAATGAACCTGTAAGTGTACCTTCCCTGACCTGAG	1444	
Db	1458	AACACCCCTGCCTGCTCGGGAAGTATTCAGGGGAGAAATTCCTGTGAACATGAAGAGATG	1517	
QY	1505	AATCACACCTGTCCTTAGAATTCCT	1530	
Db	1518	AATCACACCTGTCCTTAGAATTCCT	1543	
RESULT 2				
LOCUS	AC126895			
DEFINITION	AC126895	137964 bp	DNA linear HTG 24-JUL-2002	
ACCESSION	AC126895.1	GI:21724040		
VERSION	HTG; HTGS_PHASE1.			
KEYWORDS	Rattus norvegicus.			
SOURCE	Rattus norvegicus			
ORGANISM	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
REFERENCE				
AUTHORS	1 (bases 1 to 137964)			
	Muzny,D.M., Adams,C., Adlo-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbarella,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowler,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhal,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhardt,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Huliyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseded,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokewo,S., Oguh,M., Okunodu,G., Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubenkan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,			

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 137964)
Worley,K.C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 137964)
Worley,K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: CH230-301E4
Center clone name: CH230-301E4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89474 bases at least Q40
Consensus quality: 93422 bases at least Q30
Consensus quality: 96506 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1057 1056: contig of 1056 bp in length
1157 1156: gap of unknown length
2335 2335: contig of 1179 bp in length
2336 2435: gap of unknown length
2436 3442: contig of 1007 bp in length
3443 3542: gap of unknown length
3543 5081: contig of 1539 bp in length
5082 5181: gap of unknown length
5182 6424: contig of 1243 bp in length
6425 6524: gap of unknown length
6525 7814: contig of 1290 bp in length
7815 7914: gap of unknown length
7915 9153: contig of 1239 bp in length
9154 9253: gap of unknown length
9254 10445: contig of 1192 bp in length
10446 10545: gap of unknown length
10546 11771: contig of 1226 bp in length
11772 11871: gap of unknown length
11872 13240: contig of 1369 bp in length
13241 13340: gap of unknown length
13341 14359: contig of 1019 bp in length
14360 14459: gap of unknown length
14460 15651: contig of 1192 bp in length
15652 15751: gap of unknown length
15752 17494: contig of 1743 bp in length
17495 17594: gap of unknown length
17595 18679: contig of 1085 bp in length
18680 18779: gap of unknown length

18780 20681: contig of 1902 bp in length
20682 20781: gap of unknown length
20782 22118: contig of 1337 bp in length
22119 22218: gap of unknown length
22219 23578: contig of 1360 bp in length
23579 23678: gap of unknown length
23679 25423: contig of 1745 bp in length
25424 25523: gap of unknown length
25524 27808: contig of 2285 bp in length
27809 27908: gap of unknown length
27909 30272: contig of 2364 bp in length
30273 30372: gap of unknown length
30373 33091: contig of 2719 bp in length
33092 33191: gap of unknown length
33192 35777: contig of 2586 bp in length
35778 35877: gap of unknown length
35878 36908: contig of 1031 bp in length
36909 37008: gap of unknown length
37009 38930: contig of 1922 bp in length
38931 39030: gap of unknown length
39031 41210: contig of 2180 bp in length
41211 41310: gap of unknown length
41311 42937: contig of 1627 bp in length
42938 43037: gap of unknown length
43038 46120: contig of 3083 bp in length
46121 46220: gap of unknown length
46221 48217: contig of 1997 bp in length
48218 48317: gap of unknown length
48318 50433: contig of 2116 bp in length
50434 50533: gap of unknown length
50534 52880: contig of 2347 bp in length
52881 52980: gap of unknown length
52981 55194: contig of 2214 bp in length
55195 55294: gap of unknown length
55295 57482: contig of 2188 bp in length
57483 57582: gap of unknown length
57583 60781: contig of 3199 bp in length
60782 60881: gap of unknown length
60882 62599: contig of 1718 bp in length
62600 62699: gap of unknown length
62700 65203: contig of 2504 bp in length
65204 65303: gap of unknown length
65304 66914: contig of 1611 bp in length
66915 67015: gap of unknown length
67015 70839: contig of 3825 bp in length
70840 70939: gap of unknown length
70940 75139: contig of 4200 bp in length
75140 75239: gap of unknown length
75240 78296: contig of 3057 bp in length
78297 78396: gap of unknown length
78397 83138: contig of 4742 bp in length
83139 83238: gap of unknown length
83239 88204: contig of 4966 bp in length
88205 88304: gap of unknown length
88305 92238: contig of 3934 bp in length
92239 92338: gap of unknown length
92339 97339: contig of 5001 bp in length
97340 97439: gap of unknown length
97440 103534: contig of 6095 bp in length
103535 103634: gap of unknown length
103635 107080: contig of 3446 bp in length
107081 107180: gap of unknown length
107181 115201: contig of 8021 bp in length
115202 115301: gap of unknown length
115302 121237: contig of 5936 bp in length
121238 121337: gap of unknown length
121338 127910: contig of 6573 bp in length
127911 128010: gap of unknown length
128011 137964: contig of 9954 bp in length.

FEATURES
Source
1..137964
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"

* 6682 7773: contig of 1092 bp in length
* 7774 7873: gap of unknown length
* 7874 9530: contig of 1657 bp in length
* 9531 9630: gap of unknown length
* 9631 10661: contig of 1031 bp in length
* 10662 10761: gap of unknown length
* 10762 11958: contig of 1197 bp in length
* 11959 12058: gap of unknown length
* 12059 13140: contig of 1082 bp in length
* 13141 13240: gap of unknown length
* 13241 14777: contig of 1537 bp in length
* 14778 14877: gap of unknown length
* 14878 16063: contig of 1186 bp in length
* 16064 16163: gap of unknown length
* 16164 17585: contig of 1422 bp in length
* 17586 17685: gap of unknown length
* 17686 18916: contig of 1231 bp in length
* 18917 19016: gap of unknown length
* 19017 20305: contig of 1289 bp in length
* 20306 20405: gap of unknown length
* 20406 21537: contig of 1132 bp in length
* 21538 21637: gap of unknown length
* 21638 23703: contig of 2066 bp in length
* 23704 23803: gap of unknown length
* 23804 26306: contig of 2503 bp in length
* 26307 26406: gap of unknown length
* 26407 28431: contig of 2025 bp in length
* 28432 28531: gap of unknown length
* 28532 30615: contig of 2084 bp in length
* 30616 30715: gap of unknown length
* 30716 33215: contig of 2500 bp in length
* 33216 33315: gap of unknown length
* 33316 35357: contig of 2042 bp in length
* 35358 35457: gap of unknown length
* 35458 38012: contig of 2555 bp in length
* 38013 38112: gap of unknown length
* 38113 40338: contig of 2226 bp in length
* 40339 40438: gap of unknown length
* 40439 44051: contig of 3613 bp in length
* 44052 44151: gap of unknown length
* 44152 46294: contig of 2143 bp in length
* 46295 46394: gap of unknown length
* 46395 48845: contig of 2451 bp in length
* 48846 48945: gap of unknown length
* 48946 51753: contig of 2808 bp in length
* 51754 51853: gap of unknown length
* 51854 55915: contig of 4062 bp in length
* 55916 56015: gap of unknown length
* 56016 60468: contig of 4453 bp in length
* 60469 60568: gap of unknown length
* 60569 63640: contig of 3072 bp in length
* 63641 63740: gap of unknown length
* 63741 67888: contig of 4148 bp in length
* 67889 67988: gap of unknown length
* 67989 73014: contig of 5026 bp in length
* 73015 73114: gap of unknown length
* 73115 77313: contig of 4199 bp in length
* 77314 77414: gap of unknown length
* 77414 83786: contig of 6373 bp in length
* 83787 83886: gap of unknown length
* 83887 89590: contig of 5704 bp in length
* 89591 89690: gap of unknown length
* 89691 93846: contig of 4156 bp in length
* 93847 93946: gap of unknown length
* 93947 98126: contig of 4180 bp in length
* 98127 98226: gap of unknown length
* 98227 104258: contig of 6032 bp in length
* 104259 104358: gap of unknown length
* 104359 110994: contig of 6636 bp in length
* 110995 111094: gap of unknown length
* 111095 119838: contig of 8744 bp in length
* 119839 119938: gap of unknown length
* 119939 126339: contig of 6401 bp in length

* 126340 126439: gap of unknown length
* 126440 134888: contig of 8449 bp in length
* 134889 134988: gap of unknown length
* 134989 142709: contig of 7721 bp in length
* 142710 142809: gap of unknown length
* 142810 149818: contig of 7009 bp in length
* 149819 149918: gap of unknown length
* 149919 158719: contig of 8801 bp in length
* 158720 158819: gap of unknown length
* 158820 169868: contig of 11049 bp in length
* 169869 169968: gap of unknown length
* 169969 180390: contig of 10422 bp in length
* 180391 180490: gap of unknown length
* 180491 191272: contig of 10782 bp in length
* 191273 191372: gap of unknown length
* 191373 209667: contig of 18295 bp in length
* 209668 209767: gap of unknown length
* 209768 226303: contig of 16536 bp in length.

FEATURES
source 1. 226303
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-7C10"

BASE COUNT 57875 a 53353 c 52917 g 57041 t 5117 others

Query Match 21.6%; Score 369; DB 2; Length 226303;
Best Local Similarity 100.0%; Pred. No. 5.8e-192;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 779 AGGACACCAACTTTCTGCCCCCTTGTCACCTCCTGCTGTGATGCTCCTGTGTCAAATTG 838
Db 112254 AGGACACCAACTTTCTGCCCCCTTGTCACCTCCTGCTGTGATGCTCCTGTGTCAAATTG 112313
OY 839 GAGCAGAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCGAGTGC 898
Db 112314 GAGCAGAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCGAGTGC 112373
OY 899 CAACCAGAACCGATTCTGCCAAGTGAATGCCAACGCCGCTGTGCTGCCAGACCTG 958
Db 112374 CAACCAGAACCGATTCTGCCAAGTGAATGCCAACGCCGCTGTGCTGCCAGACCTG 112433
OY 959 CCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTAAGGCCCACTGGGGATGCG 1018
Db 112434 CCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTAAGGCCCACTGGGGATGCG 112493
OY 1019 GATACAGGGCCCTGCATCCTCAACAAATGACCCCTAGGACGAGCCCTGAGTGTGTAG 1078
Db 112494 GATACAGGGCCCTGCATCCTCAACAAATGACCCCTAGGACGAGCCCTGAGTGTGTAG 112553
OY 1079 ATGCTCTTCTCCATGCTCTTGCGCTGACGTTAACTGTCTGTGATTCACTGTGTAGAG 1138
Db 112554 ATGCTCTTCTCCATGCTCTTGCGCTGACGTTAACTGTCTGTGATTCACTGTGTAGAG 112613
OY 1139 CCACGTGAGC 1147
Db 112614 CCACGTGAGC 112622

RESULT 4
AC126895/c 137964 bp DNA linear HTG 24-JUL-2002
LOCUS AC126895 Rattus norvegicus clone CH230-301E4, *** SEQUENCING IN PROGRESS
DEFINITION *** 49 unordered pieces.
ACCESSION AC126895
VERSION AC126895.1 GI:21724040
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
REFERENCE 1 (bases 1 to 137964)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowler,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hoques,M., Hollaway,C., Hollins,B.,
Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssohn,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,S., Oguh,M., Okwuonu,G.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Nguyen,A., Nguyen,N.,
Oragunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.D., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonatke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 137964)
Worley,K.C.

Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 137964)
Worley,K.C.

Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Project name: G2HG
Center project name: CH230-301E4
Center clone name: Summary Statistics

Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89474 bases at least Q40
Consensus quality: 93422 bases at least Q30
Consensus quality: 96506 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1056: contig of 1056 bp in length
1057 1156: gap of unknown length
1157 2335: contig of 1179 bp in length
2336 2435: gap of unknown length
2436 3442: contig of 1007 bp in length
3443 3542: gap of unknown length
3543 5081: contig of 1539 bp in length
5082 5181: gap of unknown length
5182 6424: contig of 1243 bp in length
6425 6524: gap of unknown length
6525 7814: contig of 1290 bp in length
7815 7914: gap of unknown length
7915 9153: contig of 1239 bp in length
9154 9253: gap of unknown length
9254 10445: contig of 1192 bp in length
10446 10545: gap of unknown length
10546 11771: contig of 1226 bp in length
11772 11871: gap of unknown length
11872 13240: contig of 1369 bp in length
13241 13340: gap of unknown length
13341 14359: contig of 1019 bp in length
14360 14459: gap of unknown length
14460 15651: contig of 1192 bp in length
15652 15751: gap of unknown length
15752 17494: contig of 1743 bp in length
17495 17594: gap of unknown length
17595 18679: contig of 1085 bp in length
18680 18779: gap of unknown length
18780 20681: contig of 1902 bp in length
20682 20781: gap of unknown length
20782 22118: contig of 1337 bp in length
22119 22218: gap of unknown length
22219 23578: contig of 1360 bp in length
23579 23678: gap of unknown length
23679 25423: contig of 1745 bp in length
25424 25523: gap of unknown length
25524 27808: contig of 2285 bp in length
27809 27908: gap of unknown length
27909 30272: contig of 2364 bp in length
30273 30372: gap of unknown length
30373 33091: contig of 2719 bp in length
33092 33191: gap of unknown length
33192 35777: contig of 2586 bp in length
35778 35877: gap of unknown length
35879 36908: contig of 1031 bp in length
36909 37008: gap of unknown length
37009 38930: contig of 1922 bp in length
38931 39030: gap of unknown length
39031 41210: contig of 2180 bp in length
41211 41310: gap of unknown length
41311 42937: contig of 1627 bp in length
42938 43037: gap of unknown length
43038 46120: contig of 3083 bp in length
46121 46220: gap of unknown length
46221 48217: contig of 1997 bp in length
48218 48317: gap of unknown length
48318 50433: contig of 2116 bp in length
50434 50533: gap of unknown length
50534 52880: contig of 2347 bp in length
52881 52980: gap of unknown length
52981 55194: contig of 2214 bp in length
55195 55294: gap of unknown length
55295 57482: contig of 2188 bp in length
57483 57582: gap of unknown length
57583 60781: contig of 3199 bp in length
60782 60881: gap of unknown length
60882 62599: contig of 1718 bp in length

* 62600 62699: gap of unknown length
* 62700 65203: contig of 2504 bp in length
* 65204 65303: gap of unknown length
* 65304 66914: contig of 1611 bp in length
* 66915 67014: gap of unknown length
* 67015 70839: contig of 3825 bp in length
* 70840 70939: gap of unknown length
* 70940 75139: contig of 4200 bp in length
* 75140 75239: gap of unknown length
* 75240 78296: contig of 3057 bp in length
* 78297 78396: gap of unknown length
* 78397 83138: contig of 4742 bp in length
* 83139 83238: gap of unknown length
* 83239 88204: contig of 4966 bp in length
* 88205 88304: gap of unknown length
* 88305 92238: contig of 3934 bp in length
* 92239 92339: gap of unknown length
* 92339 97340 97439: gap of unknown length
* 97440 103534: contig of 6095 bp in length
* 103535 103634: gap of unknown length
* 103635 107080: contig of 3446 bp in length
* 107081 107180: gap of unknown length
* 107181 115201: contig of 8021 bp in length
* 115202 115301: gap of unknown length
* 115302 121237: contig of 5936 bp in length
* 121238 121337: gap of unknown length
* 121338 127910: contig of 6573 bp in length
* 127911 128011 137964: gap of unknown length
* 128011 137964: contig of 9954 bp in length.

FEATURES
source 1. .137964
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-301E4"
BASE COUNT 31831 a 34027 c 35229 g 31979 t 4898 others
ORIGIN

Query Match 9.7%; Score 166; DB 2; Length 137964;
Best Local Similarity 99.5%; Pred. No. 7e-80;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACGCTTCTGATCTCCAGAGACCCCTGGGGTGGGACAGGGCCCTTGCAAGGCTGCAGCC 60
Db 74321 GACGCTTCTGATCTCCAGAGACCCCTGGGGTGGGACAGGGCCCTTGCAAGGCTGCAGCC 74262
QY 61 GCTGGCAGTGGCTTGAATGAGAGTCTTTATTAATGGAAGTGAAGAGCTAAGAGGCTC 120
Db 74261 GCTGGCAGTGGCTTGAATGAGAGTCTTTATTAATGGAAGTGAAGAGCTAAGAGGCTC 74202
QY 121 CTGTCAGCTGTCTCTAAAGTCTTAGCAGCTGTGGTGGCTTGCCACACACTGTCAGA 180
Db 74201 CTGTCAGCTGTCTCTAAAGTCTTAGCAGCTGTGGTGGCTTGCCACACACTGTCAGA 74142
QY 181 CACCTTCGTGGTGGCTCCACGGCCTCACCCTTCAGGT 217
Db 74141 CACCTTCGTGGTGGCTCCACGGCCTCACCCTTCAGGT 74105

RESULT 5
AR210324 1734 bp DNA linear PAT 20-JUN-2002
LOCUS AR210324 Sequence 17 from patent US 6387657.
ACCESSION AR210324
VERSION AR210324.1 GI:21512525
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same

JOURNAL Patent: US 6387657-A 17 14-MAY-2002;
FEATURES Location/Qualifiers
source 1. .1734
BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN

Query Match 5.3%; Score 90; DB 6; Length 1734;
Best Local Similarity 100.0%; Pred. No. 1.1e-37;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGCTGCGACCCAGCCA 469
Db 418 AGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGCTGCGACCCAGCCA 477
QY 470 GGGCCTGTTGTCAAGCCTGGGGCAGGCC 499
Db 478 GGGCCTGTTGTCAAGCCTGGGGCAGGCC 507

RESULT 6
AR210325/c 1734 bp DNA linear PAT 20-JUN-2002
LOCUS AR210325 Sequence 18 from patent US 6387657.
ACCESSION AR210325
VERSION AR210325.1 GI:21512526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 18 14-MAY-2002;
FEATURES Location/Qualifiers
source 1. .1734
/organism="unknown"
BASE COUNT 393 a 495 c 491 g 355 t
ORIGIN

Query Match 5.3%; Score 90; DB 6; Length 1734;

Best Local Similarity 100.0%; Pred. No. 1.1e-37;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGCTGCGACCCAGCCA 469
Db 1317 AGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGCTGCGACCCAGCCA 1258
QY 470 GGGCCTGTTGTCAAGCCTGGGGCAGGCC 499
Db 1257 GGGCCTGTTGTCAAGCCTGGGGCAGGCC 1228

RESULT 7
AF100778 1734 bp mRNA linear ROD 17-DEC-1998
LOCUS AF100778 Mus musculus connective tissue growth factor related protein WISP-2
DEFINITION (Wisp2) mRNA, complete cds.
ACCESSION AF100778
VERSION AF100778.1 GI:4028578
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A.,
Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C.,
Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D.,
Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.
TITLE WISP genes are members of the connective tissue growth factor
family that are up-regulated in wnt-1-transformed cells and

aberrantly expressed in human colon tumors
Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)
MEDLINE 99061933
PUBMED 9843955
REFERENCE 2 (bases 1 to 1734)
AUTHORS Pennica, D.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA
Way, South San Francisco, CA 94080, USA
FEATURES
source
1. 1734
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="C57MG"
/cell_type="epithelial"
/tissue_type="mammary"
/note="transformed by Wnt-1"
1. 1734
/gene="Wisp2"
257. 1012
/gene="Wisp2"
/codon_start=1
/product="connective tissue growth factor related protein
WISP-2"
/protein_id="AAC96320.1"
/db_xref="GI:4028579"
/translation="MRGNPLIHLAISFLCISWVYSOLCPAPCACPWPPOCPGVP
LVLDGCGCCRVCAARRIGESCDLHVCDPSGGIVCOPGAGPSGRGAVCLFEEDGSCEV
NGRRYLDGETFKPNCRVLCRCDGDFGLPLGSEDEVRLPSWDCPRPRRIQVGRCCPE
WYCDQAVMOPAIOPSSAOGHOLSLVTPASADGSPCPNWSWTAMGPCSTTGLGIATRV
NONFCOLEIQRLCISRPCLASRSWSNSAP"
BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN
Query Match 5.3%; Score 90; DB 10; Length 1734;
Best Local Similarity 100.0%; Pred. No. 1.1e-37;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 410 AGTGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGTCTGCGACCCAGCCA 469
|||||
DB 418 AGTGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGTCTGCGACCCAGCCA 477
OY 470 GGGCTGTGTTGTCAGCCTGGGGCAGGCC 499
|||||
DB 478 GGGCTGTGTTGTCAGCCTGGGGCAGGCC 507
RESULT 8
AL731698 61072 bp DNA linear ROD 24-MAY-2002
LOCUS AL731698
DEFINITION Mouse DNA sequence from clone RP23-161B3 on chromosome 2, complete
sequence.
ACCESSION AL731698
VERSION AL731698.10 GI:21214309
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 61072)
Wallis, J.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 25, 2002 this sequence version replaced gi:21213601.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, ask we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-161B3 is
from the RPI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6.
FEATURES
source
1. 61072
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-161B3"
/clone_lib="RPI-23"
BASE COUNT 15405 a 15368 c 15033 g 15266 t
ORIGIN
Query Match 5.3%; Score 90; DB 10; Length 61072;
Best Local Similarity 100.0%; Pred. No. 7e-38;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 410 AGTGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGTCTGCGACCCAGCCA 469
|||||
DB 10029 AGTGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGTCTGCGACCCAGCCA 10088
OY 470 GGGCTGTGTTGTCAGCCTGGGGCAGGCC 499
|||||
DB 10089 GGGCTGTGTTGTCAGCCTGGGGCAGGCC 10118
RESULT 9
AL669906 216757 bp DNA linear HTG 24-JUL-2002
LOCUS AL669906
DEFINITION Mus musculus chromosome 2 clone RP23-217C2, *** SEQUENCING IN
PROGRESS ***; 21 unordered pieces.
ACCESSION AL669906 GI:21955520
VERSION AL669906.5
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 216757)
Sims, S.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 25, 2002 this sequence version replaced gi:18181793.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: bm217C2
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 211865 bases at least Q40
Consensus quality: 213403 bases at least Q30
Consensus quality: 214139 bases at least Q20
Insert size: 214757; sum-of-contigs
Insert size: 234243; 1.8% error; agarose-fp
Quality coverage: 6.29x in Q20 bases; sum-of-contigs Quality

Coverage: 5.90x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 4000: contig of 4000 bp in length
* 4001 4100: gap of 100 bp
* 4101 15267: contig of 11167 bp in length
* 15268 15367: gap of 100 bp
* 15368 34970: contig of 19603 bp in length
* 34971 35070: gap of 100 bp
* 35071 39671: contig of 4601 bp in length
* 39672 39771: gap of 100 bp
* 39772 46360: contig of 6589 bp in length
* 46361 46460: gap of 100 bp
* 46461 49809: contig of 3349 bp in length
* 49810 49909: gap of 100 bp
* 49910 54509: contig of 4600 bp in length
* 54510 54609: gap of 100 bp
* 54610 65989: contig of 11380 bp in length
* 65990 66089: gap of 100 bp
* 66090 84635: contig of 18546 bp in length
* 84636 84735: gap of 100 bp
* 84736 87399: contig of 2664 bp in length
* 87400 87499: gap of 100 bp
* 87500 98601: contig of 11102 bp in length
* 98602 98701: gap of 100 bp
* 98702 105522: contig of 6821 bp in length
* 105523 105622: gap of 100 bp
* 105623 119773: contig of 14151 bp in length
* 119774 119873: gap of 100 bp
* 119874 134552: contig of 14679 bp in length
* 134553 134652: gap of 100 bp
* 134653 141202: contig of 6550 bp in length
* 141203 141302: gap of 100 bp
* 141303 145808: contig of 4506 bp in length
* 145809 145908: gap of 100 bp
* 145909 151572: contig of 5664 bp in length
* 151573 151672: gap of 100 bp
* 151673 169686: contig of 18014 bp in length
* 169687 169786: gap of 100 bp
* 169787 200095: contig of 30309 bp in length
* 200096 200195: gap of 100 bp
* 200196 213877: contig of 13682 bp in length
* 213878 213977: gap of 100 bp
* 213978 216757: contig of 2780 bp in length.

FEATURES
Source Location/Qualifiers

1. .216757
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-217C2"
/clone_lib="RPCT-23"
1. .4000
/note="assembly_fragment:03429
clone_end:SP6
vector_side:left"
4101. .15267
/note="assembly_fragment:00814
fragment_chain:1"
15368. .34970
/note="assembly_fragment:00812
fragment_chain:1"
35071. .39671
/note="assembly_fragment:03544
fragment_chain:1"
39772. .46360
/note="assembly_fragment:01438

misc_feature fragment_chain:1"
46461. .49809
/note="assembly_fragment:03410
fragment_chain:2"
49910. .54509
/note="assembly_fragment:00811
fragment_chain:2"
54610. .65989
/note="assembly_fragment:02217
fragment_chain:2"
66090. .84635
/note="assembly_fragment:03529
fragment_chain:2"
84736. .87399
/note="assembly_fragment:01019
fragment_chain:3"
87500. .98601
/note="assembly_fragment:01170
fragment_chain:3"
98702. .105522
/note="assembly_fragment:02829
fragment_chain:3"
105623. .119773
/note="assembly_fragment:01476
fragment_chain:4"
119874. .134552
/note="assembly_fragment:00505
fragment_chain:4"
134653. .141202
/note="assembly_fragment:01951
fragment_chain:4"
141303. .145808
/note="assembly_fragment:02637
fragment_chain:5"
145909. .151572
/note="assembly_fragment:02352
fragment_chain:5"
151673. .169686
/note="assembly_fragment:02207
fragment_chain:5"
169787. .200095
/note="assembly_fragment:02849
fragment_chain:6"
200196. .213877
/note="assembly_fragment:01052
fragment_chain:6"
213978. .216757
/note="assembly_fragment:01318
clone_end:T7
vector_side:right"
BASE COUNT 56425 a 51573 c 51904 g 54852 t 2003 others
ORIGIN

Query Match 5.3%; Score 90; DB 2; Length 216757;
Best Local Similarity 100.0%; Pred. No. 5.9e-38;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTCACGAGGCTGGGGAGTCTGCACACCTGCATGTCGACCCAGCCA 469
|||||
Db 166450 AGTGTGTCACGAGGCTGGGGAGTCTGCACACCTGCATGTCGACCCAGCCA 166509
QY 470 GGGCCTGTTGTCTCAGCCTGGGGCAGGCC 499
|||||
Db 166510 GGGCCTGTTGTCTCAGCCTGGGGCAGGCC 166539

RESULT 10
AF126063 1739 bp mRNA linear ROD 12-OCT-1999
LOCUS AF126063
DEFINITION Mus musculus connective tissue growth factor-like protein precursor
(Ctgf1) mRNA, complete cds.
ACCESSION AF126063
VERSION AF126063.1 GI:4337059

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Mus musculus.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1739)
Kumar, S., Hand, A.T., Connor, J.R., Dodds, R.A., Ryan, P.J., Trill, J.J., Fisher, S.M., Nuttall, M.E., Lipschutz, D.B., Zou, C., Hwang, S.M., Volta, B.J., James, I.E., Riemann, D.J., Gowen, M. and Lee, J.C.
Identification and cloning of a connective tissue growth factor-like cDNA from human osteoblasts encoding a novel regulator of osteoblast functions
J. Biol. Chem. 274 (24), 17123-17131 (1999)
99287915
10358067
2 (bases 1 to 1739)
Kumar, S. and Zou, C.
Direct Submission
Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109, SmithKline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406, USA

FEATURES
source
1. .1739
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="Lung"
1. .1739
/gene="Ctgfl"
242. .997
/gene="Ctgfl"
/note="similar to the Mus musculus WISP-2 protein encoded by the sequence presented in GenBank Accession Number AF100778; putative growth factor; CTGF-L; contains IGF binding (IGFBP), Von Willibrand Factor type C (VWC) repeat and thrombospondin type I (TSP1) domains; member of the CCN (CTGF/Cyrl61/Nov) family; lacks the fourth carboxy-terminal (CT) domain present in other members of the CCN family"
/codon_start=1
/product="connective tissue growth factor-like protein precursor"
/protein_id="AA018058.1"
/db_xref="GI:4337060"
/translation="MRGNPLIHLLAIFLCILSMVYSQCPAPACFPTPPQCPGPV LVLGGCCRCRYCARLIGESCDHLHVCDSQGLVCQPGAGPSGRVCLFEEDGSCVE NGRRLDGETFEKPNCRVLCRCDDGGFTCLPFCSEDRVLPWDCPPRRIOVPGRCPE WWCDAVMQPAIQSSAQGHLSALVTPASADGPCPNWSTAMGPCSTTCGLIATRVSNONRCQLEIQRLCLSRPCLASRSHGSWNSAF"

BASE COUNT
ORIGIN
375 a 480 c 489 g 395 t

Query Match
Best Local Similarity 100.0%; Score 80; DB 10; Length 1739;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 CGGAGGCTGGGGAGTCTCGGACACCTGCATGCTCGGACCCAGCCAGGCGCTGGTT 479
|||||
Db 413 CGGAGGCTGGGGAGTCTCGGACACCTGCATGCTCGGACCCAGCCAGGCGCTGGTT 472
|||||

QY 480 TGTACAGCTGGGGCAGGCC 499
|||||
Db 473 TGTACAGCTGGGGCAGGCC 492
|||||

RESULT 11
AR210337 738 bp DNA linear PAT 20-JUN-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
Sequence 38 from patent US 6387657.
AR210337
AR210337.1 GI:21512542
Unknown.

ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 32; DB 6; Length 738;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGTGCTGATGCTGTGCTGTG 406
|||||
Db 115 GTACCCCTGTGCTGATGCTGTGCTGTG 146
|||||

RESULT 12
AR210338 841 bp DNA linear PAT 20-JUN-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 32; DB 6; Length 841;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGATGTGCGGCTGCCAGCTGGGACTGCC 685
|||||
Db 417 GAGATGTGCGGCTGCCAGCTGGGACTGCC 448
|||||

RESULT 13
AX076919 1266 bp DNA linear PAT 22-FEB-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
Sequence 31 from patent WO0105836.
AX076919
AX076919.1 GI:13121575
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1266)
Botstein, D., Goddard, A., Gurney, A.L., Hillan, K.J., Roy, M.A. and Wood, W.I.
Polypeptidic compositions and methods for the treatment of tumors
Patent: WO 0105836-A 31 25-JAN-2001;
Genentech, Inc. (US)
Location/Qualifiers
1. .1266
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 216 a 418 c 390 g 242 t
ORIGIN
Query Match 1.9%; Score 32; DB 6; Length 1266;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 375 GTACCCCTGGTCTGATGGCTGTGGCTGCTG 406
Db 136 GTACCCCTGGTCTGATGGCTGTGGCTGCTG 167
RESULT 14
AX464186 1266 bp DNA linear PAT 16-JUL-2002
LOCUS AX464186
DEFINITION Sequence 319 from Patent WO0140466.
ACCESSION AX464186
VERSION AX464186.1 GI:21899109
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,
AUTHORS Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
Wood,W.L. and Zhang,Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
same
JOURNAL Patent: WO 0140466-A 319 07-JUN-2001;
Genentech Inc. (US)
FEATURES Location/Qualifiers
source 1..1266
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 216 a 418 c 390 g 242 t
ORIGIN
Query Match 1.9%; Score 32; DB 6; Length 1266;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 375 GTACCCCTGGTCTGATGGCTGTGGCTGCTG 406
Db 136 GTACCCCTGGTCTGATGGCTGTGGCTGCTG 167
RESULT 15
AF083500 1283 bp mRNA linear PRI 04-NOV-1998
LOCUS AF083500
DEFINITION Homo sapiens connective tissue growth factor-like protein
precursor, mRNA, complete cds.
ACCESSION AF083500
VERSION AF083500.1 GI:3462835
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1283)
AUTHORS Kumar,S., Hand,A.T., Connor,J.C., Dodds,R.A., Ryan,P.J.,
Trill,J.G., Fisher,S.M., Slemmon,J.R., Lipshutz,D.B.,
Bartholomew,V., James,I.E., Rieman,D.J., Gowen,M. and Lee,J.C.
TITLE Identification and cloning of CTGF-L from human osteoblasts, a
novel cysteine rich protein containing an IGF binding domain
JOURNAL Bone 23 (5), S240 (1998)
REFERENCE 2 (bases 1 to 1283)
AUTHORS Kumar,S.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1998) Bone & Cartilage Biology, UW 2109,
SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406,
USA

FEATURES Location/Qualifiers
source 1..1283
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q12-q13"
/cell_type="primary osteoblast"
9..761
/note="CTGF-L; encodes IGF binding (IGFB), Von Willebrand
factor type C (WVC) and thrombospondin type I (TSP1)
domains; member of the CCN (CTGF/cyrl61/nov) family; lacks
the fourth carboxy-terminal domain present in other
members of the CCN family"
/codon_start=1
/product="connective tissue growth factor-like protein
precursor"
/protein_id="AAC70350.1"
/db_xref="GI:3462836"
/translation="MRGTPKTHLLAFSLCLLSKVRTQLCPPTCPWPPPRCPPLGVP
LVLDGGCCRCVCA RLGEPCDQLHYCDASQGLVCCPGAGPGRGALCLLAEDSSCEV
NGRLYREGETFQPHCSIRCRCEDEGFTCVPLCSEDEVRLPSMDCPPHPRVEVLGKCCE
WVCGGGGLGTQPLPAQGPQFSGLVSSLPPGVPCPEWSTAWGPCSTTCGLGMATRVSN
QNRFCRIETQRRICLSRPPCPSPRGSRSPQNSAF"
BASE COUNT 235 a 418 c 389 g 241 t
ORIGIN
Query Match 1.9%; Score 32; DB 9; Length 1283;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 375 GTACCCCTGGTCTGATGGCTGTGGCTGCTG 406
Db 135 GTACCCCTGGTCTGATGGCTGTGGCTGCTG 166

Search completed: July 28, 2003, 21:14:39
Job time : 4755.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:31.11 ; Search time 448.586 Seconds
(without alignments)
8574.520 Million cell updates/sec

Title: US-10-010-408-1
Perfect score: 1708
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GCCCTAGATAAACACCCAAA 1708

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 segs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1708	100.0	1708	20	AAZ07516	Rat HICP polypepti
2	753	44.1	753	20	AAZ07517	Rat HICP polypepti
3	681	39.9	681	20	AAZ07521	Rat HICP mature po
4	354	20.7	439	24	ABL59575	Rat OST23 gene fra
5	210	12.3	210	20	AAZ07519	Rat HICP IGFBP dom
6	177	10.4	177	20	AAZ07518	Rat HICP VWC doma
7	174	10.2	174	20	AAZ07520	Rat HICP TSPI doma
8	90	5.3	753	20	AAZ76489	Mouse WISP-2 prote
9	90	5.3	1734	20	AAZ76488	Mouse WISP-2 prote

10	54	3.2	65	24	ABN30189	Rat spliced trans
11	32	1.9	199	22	ABA72245	Human foetal liver
12	32	1.9	199	22	ABA38112	Probe #16578 for g
13	32	1.9	199	22	AAK20667	Human brain expres
14	32	1.9	199	22	AAK46811	Human bone marrow
15	32	1.9	199	22	AAI25749	Probe #15682 for g
16	32	1.9	199	22	AAI52650	Probe #21336 used
17	32	1.9	199	24	AB521124	Human genome-deriv
18	32	1.9	586	22	ABA59703	Human foetal liver
19	32	1.9	586	22	ABA28232	Probe #6698 for ge
20	32	1.9	586	22	AAK07972	Human brain expres
21	32	1.9	586	22	AAK33842	Human bone marrow
22	32	1.9	586	22	AAI16552	Probe #6485 for ge
23	32	1.9	586	22	AAI39568	Probe #8254 used t
24	32	1.9	586	24	ABS08689	Human genome-deriv
25	32	1.9	738	20	AAZ76501	Human WISP-2 prote
26	32	1.9	841	20	AAZ76502	Human WISP-2 prote
27	32	1.9	841	20	AAZ76502	Human WISP-2 prote
28	32	1.9	1257	20	AAZ28435	Human WISP-2 prote
29	32	1.9	1266	21	AAA30048	Human PRO261 nucle
30	32	1.9	1266	22	AA521403	Human cDNA sequenc
31	32	1.9	1266	22	AAF60368	PRO261 coding sequ
32	32	1.9	1266	22	AAC97451	Human angiogenesis
33	32	1.9	1267	21	AAA77566	Human PRO261 cDNA
34	32	1.9	1285	19	AAV29260	Human connective t
35	32	1.9	1293	20	AAZ76486	Human WISP-2 prote
36	32	1.9	1309	22	AAH28214	Human secreted pro
37	32	1.9	1337	22	AAH46952	Human secreted pro
38	32	1.9	1352	22	AAH46936	Human secreted pro
39	32	1.9	1522	20	AAZ16595	Human growth facto
40	32	1.9	2136	22	AAK94706	Human full-length
41	32	1.9	2136	22	AAK94706	Human immune/haema
42	27	1.6	51	20	AAZ76535	Human WISP-2 probe
43	27	1.6	51	21	AAA76569	Human PRO261 hybri
44	27	1.6	51	21	AAA30051	Probe for PRO261 n
45	27	1.6	51	22	AAF60371	PRO261 probe #1.

ALIGNMENTS

RESULT 1	
AAZ07516	
ID AAZ07516	standard; cDNA; 1708 BP.
AC AAZ07516;	
XX	
DT 26-NOV-1999	(first entry)
XX	
DE	Rat HICP polypeptide encoding cDNA.
XX	
KW	Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW	cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX	
OS	Rattus sp.
XX	
PN	W09947556-A2.
XX	
PD	23-SEP-1999.
XX	
PF	18-MAR-1999; 99WO-US05999.
XX	
PR	19-MAR-1998; 98US-0044273.
XX	
PA	(TUFT) TUFTS COLLEGE.
XX	
PI	Castellot JJ;
XX	
DR	WPI; 1999-562060/47.
DR	P-PSDB; AAY27434.
XX	
PT	Nucleic acid sequences encoding rat heparin-induced CCN-like protein, - used in methods to identify modulators or in diagnostic applications

XX Claim 2; Fig 1; 108bp; English.
XX
CC This cDNA encodes a rat heparin-induced CN-1-like protein (HICP) protein.
CC Agents that stimulate or inhibit HICP protein activity or expression,
CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to
CC modulate cell-associated activity. HICP modulators can be used to treat
CC disorders characterized by aberrant HICP protein activity or expression.
CC Probes capable of hybridizing to HICP mRNA or antibodies specific for
CC HICP can be used to detect HICP activity in a biological sample. HICP
CC can be used to treat disorders, such as a cardiovascular or fibrotic
CC disorder, characterized by aberrant cell proliferation.
XX
SQ Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Query Match 100.0%; Score 1708; DB 20; Length 1708;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGCTTGTATCTCCAGAGGACCCTGGGTGGACAGGGGCGCTGGCAGCGTGCAGCC 60
Db 1 GACGCTTGTATCTCCAGAGGACCCTGGGTGGACAGGGGCGCTGGCAGCGTGCAGCC 60
OY 61 GCTGGGAGTGGCTTGAATGAGGTCTTTATTACTGGGAACTAGAGAGCTAAGAGGCTC 120
Db 61 GCTGGGAGTGGCTTGAATGAGGTCTTTATTACTGGGAACTAGAGAGCTAAGAGGCTC 120
OY 121 CTGTACAGCTGTCTCTAAAGTCTTAGACACTGTGTGGCTTGGGCTTCACACACTGTGAGA 180
Db 121 CTGTACAGCTGTCTCTAAAGTCTTAGACACTGTGTGGCTTGGGCTTCACACACTGTGAGA 180
OY 181 CACCTTCGTGTGCTCCACAGGCGCTCACCTTCAGGTTGAAGCTGGCTCCACAAGGAGC 240
Db 181 CACCTTCGTGTGCTCCACAGGCGCTCACCTTCAGGTTGAAGCTGGCTCCACAAGGAGC 240
OY 241 ACGGTGACATGAGGGGACAGCCCACTGATTCATCTTGGGCACTTCTCTGCTTC 300
Db 241 ACGGTGACATGAGGGGACAGCCCACTGATTCATCTTGGGCACTTCTCTGCTTC 300
OY 301 TCTCAATGCTGTGTGCTCCAGCTGTGCGGACACCTGTAGCTGTCTGAGACACACCC 360
Db 301 TCTCAATGCTGTGTGCTCCAGCTGTGCGGACACCTGTAGCTGTCTGAGACACACCC 360
OY 361 AGTCCACACAGGGGGTACCCCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 420
Db 361 AGTCCACACAGGGGGTACCCCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 420
OY 421 GGAGGCTGGGGAGTCTGCGACACCTGATGTGCGAACCCACAGCGGCGCTGTGTT 480
Db 421 GGAGGCTGGGGAGTCTGCGACACCTGATGTGCGAACCCACAGCGGCGCTGTGTT 480
OY 481 GTACGCTGGGGAGGCGCTGGCGGCAATGGGCTGTGTCTCTTGATGAGGATGACG 540
Db 481 GTACGCTGGGGAGGCGCTGGCGGCAATGGGCTGTGTCTCTTGATGAGGATGACG 540
OY 541 GTAGCTGTGAGGTGAATGGCCGAGGTACTGTGATGAGAGAGACCTTTAAACCAATTGCA 600
Db 541 GTAGCTGTGAGGTGAATGGCCGAGGTACTGTGATGAGAGAGACCTTTAAACCAATTGCA 600
OY 601 GGGTCTGTGCTGTGATGAGCGGTGCTTCACTGCGCGCTGTGCAAGTGAAGATG 660
Db 601 GGGTCTGTGCTGTGATGAGCGGTGCTTCACTGCGCGCTGTGCAAGTGAAGATG 660
OY 661 TGGGCTGGCCAGCTGGGAGCTGCCACAGGCCCAAGAGATACAGGTGCGAGGAAAGTGCT 720
Db 661 TGGGCTGGCCAGCTGGGAGCTGCCACAGGCCCAAGAGATACAGGTGCGAGGAAAGTGCT 720
OY 721 GCCCGAGTGGGTATGTGACAGGAGGTGACACCGGCGATCCAGCGCTCCACGGCGCAAG 780
Db 721 GCCCGAGTGGGTATGTGACAGGAGGTGACACCGGCGATCCAGCGCTCCACGGCGCAAG 780
OY 781 GACACCAACTTCTGCGCTGTGCTACTCTGCTCTGTGATGCTCTGTTCCAAATTGGA 840
Db 781 GACACCAACTTCTGCGCTGTGCTACTCTGCTCTGTGATGCTCTGTTCCAAATTGGA 840

Db 781 GACACCAACTTCTGCGCTGTGCTACTCTGCTCTGTGATGCTCTGTTCCAAATTGGA 840
OY 841 GCACAGCCTGGGGCCCCCTGCTCAACACCACTGTGGCTGGGCATAGCCACCGAGTGTCCA 900
Db 841 GCACAGCCTGGGGCCCCCTGCTCAACACCACTGTGGCTGGGCATAGCCACCGAGTGTCCA 900
OY 901 ACCAGAACCGATTTCTGCCAAGTGAAGATCAACCGCGCTGTGTGCTGCCAGACCTGCGC 960
Db 901 ACCAGAACCGATTTCTGCCAAGTGAAGATCAACCGCGCTGTGTGCTGCCAGACCTGCGC 960
OY 961 TGGCAGCCAGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCAAGTGGGATGCGGA 1020
Db 961 TGGCAGCCAGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCAAGTGGGATGCGGA 1020
OY 1021 TACAGGCGCTGCCATCTCTACAGCAATGACCCCTAGACACAGCGCCCTGCACTGTAGAT 1080
Db 1021 TACAGGCGCTGCCATCTCTACAGCAATGACCCCTAGACACAGCGCCCTGCACTGTAGAT 1080
OY 1081 GCTCTTCTCCATGCTCTTGGCTGCAGTTAATGTCTCTGCTTGGATTCACTGTGTAGAGCC 1140
Db 1081 GCTCTTCTCCATGCTCTTGGCTGCAGTTAATGTCTCTGCTTGGATTCACTGTGTAGAGCC 1140
OY 1141 ACTGAGCGATCCCTGCTCTGTCTGAGTAGGCGGAGCAGGTGACACAGCTCCAGTCTCTG 1200
Db 1141 ACTGAGCGATCCCTGCTCTGTCTGAGTAGGCGGAGCAGGTGACACAGCTCCAGTCTCTG 1200
OY 1201 GTTCAGCCTGGAATTTCTGGGTCTCTGCTGCTCATTTCTCAAAACATCCCTGTACAAAAG 1260
Db 1201 GTTCAGCCTGGAATTTCTGGGTCTCTGCTGCTCATTTCTCAAAACATCCCTGTACAAAAG 1260
OY 1261 GACAACCAAAAAGACCTTTAAACCTAGGCTATACCTGGCAAACTGGCCACCGTGTGG 1320
Db 1261 GACAACCAAAAAGACCTTTAAACCTAGGCTATACCTGGCAAACTGGCCACCGTGTGG 1320
OY 1321 GATAAGGTCAATGTTAGACACAGACAGAGATGCTGAAACTTCCCAATCCCTCTTGG 1380
Db 1321 GATAAGGTCAATGTTAGACACAGACAGAGATGCTGAAACTTCCCAATCCCTCTTGG 1380
OY 1381 ACTTCTGATGCTTGTCCCAAGATGATGAATGAATCTGTAAGTACCTTCCCTGACC 1440
Db 1381 ACTTCTGATGCTTGTCCCAAGATGATGAATGAATCTGTAAGTACCTTCCCTGACC 1440
OY 1441 TGAGAACACCCCTGCTGCTGGGAGATATTACAGGGGACAGAAATTTCTGTGAACATGAAGA 1500
Db 1441 TGAGAACACCCCTGCTGCTGGGAGATATTACAGGGGACAGAAATTTCTGTGAACATGAAGA 1500
OY 1501 GATGAATCACACTGTCTTAAAGAAATTCCTGAAGTCCAGGAACCTTGATTTT 1560
Db 1501 GATGAATCACACTGTCTTAAAGAAATTCCTGAAGTCCAGGAACCTTGATTTT 1560
OY 1561 CAGGAATGACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACTCTGGCAGGCC 1620
Db 1561 CAGGAATGACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACTCTGGCAGGCC 1620
OY 1621 AGGCGCTTCTCTTACGATGAGAAAGACAGGAGCAGCAGAGTACTCTCTGAGAGA 1680
Db 1621 AGGCGCTTCTCTTACGATGAGAAAGACAGGAGCAGCAGAGTACTCTCTGAGAGA 1680
OY 1681 CTAGCTAGCCTAGATAAACCACCAAA 1708
Db 1681 CTAGCTAGCCTAGATAAACCACCAAA 1708

RESULT 2
AAZ07517
ID AAZ07517 standard; cDNA; 753 BP.
XX
AC AAZ07517;
XX
DT 26-NOV-1999 (first entry)
XX
DE Rat HICP polypeptide coding sequence.
XX

KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX
OS Rattus sp.
XX
PN W09947556-A2.
XX
PD 23-SEP-1999.
XX
PE 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Castellot JU;
XX
DR WPI; 1999-562060/47.
DR P-PSDB; AAY27434.
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications
XX
PS Claim 5; Fig 1; 108pp; English.
XX
CC The invention provides a rat heparin-induced CCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents the coding sequence of rat HICP.
XX
SQ Sequence 753 BP; 132 A; 235 C; 230 G; 156 T; 0 other;

Query Match 44.1%; Score 753; DB 20; Length 753;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCCTCTGCTCTCTCAATG 308
Db 1 ATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCCTCTGCTCTCTCAATG 60
QY 309 GTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTGGACACCAACCCAGTGCCCA 368
Db 61 GTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTCTGGACACCAACCCAGTGCCCA 120
QY 369 CAGGGGGTACCCCTGTGCTGTGATGGCTGTGGCTGTGTAAGTGTGTGCACGGAGGCTG 428
Db 121 CAGGGGGTACCCCTGTGCTGTGATGGCTGTGGCTGTGTAAGTGTGTGCACGGAGGCTG 180
QY 429 GGGGAGTCTTCGCGACCACTGATGTCTGCGACCCCAAGGCGCTGTTGTACGCT 488
Db 181 GGGGAGTCTTCGCGACCACTGATGTCTGCGACCCCAAGGCGCTGTTGTACGCT 240
QY 489 GGGGAGGCGCTGCGGGCCATGGGGCTGTGTCTCTTGGATGAGATGACGCTAGCTGT 548
Db 241 GGGGAGGCGCTGCGGGCCATGGGGCTGTGTCTCTTGGATGAGATGACGCTAGCTGT 300
QY 549 GAGGTGAATGGCGCAGGTACCTGATGGAGAGACCTTTAAACCAATTGCAGGGTCTCTG 608
Db 301 GAGGTGAATGGCGCAGGTACCTGATGGAGAGACCTTTAAACCAATTGCAGGGTCTCTG 360
QY 609 TGCCGCTGTGATGACGCTGCTTCACTTGCCTGCCGCTGTGCAGTGAGATGTGCGGCTG 668
Db 361 TGCCGCTGTGATGACGCTGCTTCACTTGCCTGCCGCTGTGCAGTGAGATGTGCGGCTG 420
QY 669 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAGAAAGTGTGCCCCGAG 728
Db 421 CCCAGCTGGGACTGCCACGCCCCAAGAGATACAGGTGCCAGAAAGTGTGCCCCGAG 480

QY 729 TGGGTATGTGACCAAGGAGTGAACACGGCGATCCAGCGCTCCACAGGGCGCAAGACACCA 788
Db 481 TGGGTATGTGACCAAGGAGTGAACACGGCGATCCAGCGCTCCACAGGGCGCAAGACACCA 540
QY 789 CTTTCTGCCCTTGTCACTCTCGCTCTGCTGTATGCTCTCTGTCCAAATTTGGAGCACAGCC 848
Db 541 CTTTCTGCCCTTGTCACTCTCGCTCTGCTGTATGCTCTCTGTCCAAATTTGGAGCACAGCC 600
QY 849 TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAC 908
Db 601 TGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACCAAGAC 660
QY 909 CGATTCTGCCCACTGAGATCCACGCGCCTGTGTCTGCCAGACCCCTGCTGCGAGCC 968
Db 661 CGATTCTGCCCACTGAGATCCACGCGCCTGTGTCTGCCAGACCCCTGCTGCGAGCC 720
QY 969 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 1001
Db 721 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 753

RESULT 3

ID AAZ07521 standard; cDNA; 681 BP.
XX
AC AAZ07521;
XX
DT 26-NOV-1999 (first entry)
XX
DE Rat HICP mature polypeptide coding sequence.

KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX
OS Rattus sp.

PN W09947556-A2.
XX
PD 23-SEP-1999.
XX
PE 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Castellot JU;
XX

DR WPI; 1999-562060/47.
DR P-PSDB; AAY27440.

PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications
XX
PS Disclosure; Fig 2; 108pp; English.

CC The invention provides a rat heparin-induced CCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents the coding sequence of rat HICP mature
CC polypeptide.

SQ Sequence 681 BP; 121 A; 211 C; 215 G; 134 T; 0 other;

Query Match 39.9%; Score 681; DB 20; Length 681;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
OY	318	CAGCTGTGCGGACACCCCTGTACCTGTCTCTTGGACACCAACCCAGTGGCCACAGGGGTA	377		
Db	1	CAGCTGTGCGGACACCCCTGTACCTGTCTCTTGGACACCAACCCAGTGGCCACAGGGGTA	60		
OY	378	CCCCGTGTGTGGATGGCTGTGGCTGTCTTAAAGTGTGTGCACGGAGGCTGGGGAGTCC	437		
Db	61	CCCCGTGTGTGGATGGCTGTGGCTGTCTTAAAGTGTGTGCACGGAGGCTGGGGAGTCC	120		
OY	438	TGCGACCACTGCATGTCTGCGACCCACGCGGCGCTGTGTTGTCAAGCTGGGGCAGGC	497		
Db	121	TGCGACCACTGCATGTCTGCGACCCACGCGGCGCTGTGTTGTCAAGCTGGGGCAGGC	180		
OY	498	CCTGGCGGCAATGGGCTGTGTCTCTTGGATGAGATGACGCTAGCTGTGAGTGAAT	557		
Db	181	CCTGGCGGCAATGGGCTGTGTCTCTTGGATGAGATGACGCTAGCTGTGAGTGAAT	240		
OY	558	GGCCGACGTAACCTGGATGGAGAGACCTTTAAACCAATGTCAGGGTCTGTCGGCTGT	617		
Db	241	GGCCGACGTAACCTGGATGGAGAGACCTTTAAACCAATGTCAGGGTCTGTCGGCTGT	300		
OY	618	GATGACGGTGGCTTCAACCTGCCCTGCCCTGTGCAGTGAAGATGTGGCGTCCCAAGCTGG	677		
Db	301	GATGACGGTGGCTTCAACCTGCCCTGCCCTGTGCAGTGAAGATGTGGCGTCCCAAGCTGG	360		
OY	678	GACTGCCCCACGCCCCAAGAGATACAGGTGCGCAGGAAAGTGTGCCCGCAGTGGGTATGT	737		
Db	361	GACTGCCCCACGCCCCAAGAGATACAGGTGCGCAGGAAAGTGTGCCCGCAGTGGGTATGT	420		
OY	738	GACCAGGAGTGACACCGCGGATCCAGCGCTCCACGGCGCAAGAGACCACTTCTGCGC	797		
Db	421	GACCAGGAGTGACACCGCGGATCCAGCGCTCCACGGCGCAAGAGACCACTTCTGCGC	480		
OY	798	CTTGTCACTCTGCTCTGTGTGATGCTCTTGTCCAATGTGAGACACAGCTGGGGCCCC	857		
Db	481	CTTGTCACTCTGCTCTGTGTGATGCTCTTGTCCAATGTGAGACACAGCTGGGGCCCC	540		
OY	858	TGCTCAACCACTGTGGGCTGGCATAGCCACCCGAGTGTCCAACCAAGACCATTTCTGC	917		
Db	541	TGCTCAACCACTGTGGGCTGGCATAGCCACCCGAGTGTCCAACCAAGACCATTTCTGC	600		
OY	918	CAACTGGAGATCAACGCGGCTGTGTCTGTGCCAGACCCCTGGCTGGCAGCAGAGCCAC	977		
Db	601	CAACTGGAGATCAACGCGGCTGTGTCTGTGCCAGACCCCTGGCTGGCAGCAGAGCCAC	660		
OY	978	AGCTCATGGAACAGTGTCTTC 998			
Db	661	AGCTCATGGAACAGTGTCTTC 681			
RESULT 4					
ABL59575					
ID	ABL59575 standard; DNA; 439 BP.				
XX	ABL59575;				
XX	17-JUL-2002 (first entry)				
DE	Rat OST23 gene fragment SEQ ID NO:23.				
XX	Rat: OST; osteoregenerative; parathyroid hormone; tibiae; osteopathic;				
KW	gene therapy; parathyroid hormone receptor ligand; bone disorder;				
KW	bone formation disorder; bone resorption disorder; osteoporosis;				
KW	osteopenia; osteopetrosis; gene; ds.				
OS	Rattus sp.				
PN	W0200224943-A2.				
XX	28-MAR-2002.				
PD	19-SEP-2001; 2001WO-US29548.				
PF					

XX	PR	19-SEP-2000; 2000US-233579P.			
XX	PA	(CURA-) CURAGEN CORP.			
PA	PA	(GLAX) GLAXO GROUP LTD.			
XX	PI	Horesovsky GJ, Noel LS, Raha D;			
XX	DR	WPI; 2002-401989/43.			
PT	PT	Identifying parathyroid hormone receptor ligands and osteoregenerative			
PT	PT	agents involves detecting the expression of nucleic acids which are			
PT	PT	regulated by parathyroid hormone			
XX	PS	Claim 45; Page 22; 90pp; English.			
XX	CC	The present invention describes a method (M1) for identifying parathyroid			
CC	CC	hormone receptor ligands (I) and osteoregenerative agents by contacting a			
CC	CC	test cell population (CP) comprising cells expressing nucleic acid			
CC	CC	sequences (S) of OST1-47 and 48, with a test agent, measuring nucleic			
CC	CC	acid sequence expression, comparing it with reference CP and identifying			
CC	CC	the ligand and agent by the difference in expression levels. Also			
CC	CC	described is a method (M2) for treating a bone disorder in a subject by			
CC	CC	administering to the subject an agent that modulates the expression or			
CC	CC	activity of (S). (I) have osteopathic activities, and can be used to			
CC	CC	modulate the expression of OST 1-48, and can also be used in gene			
CC	CC	therapy. (M1) is useful for identifying parathyroid hormone receptor			
CC	CC	ligands (I) or osteoregenerative agents. OST gene sequence can be used			
CC	CC	for assessing the osteoregenerative activity of a test agent in a			
CC	CC	subject, and for diagnosing or determining the susceptibility to bone			
CC	CC	disorder and assessing efficacy of a treatment of a bone disorder in a			
CC	CC	subject, human or rodent. (M2) is useful for treating a bone disorder			
CC	CC	including bone formation disorder or bone resorption disorder (e.g.			
CC	CC	osteoporosis, osteopenia and osteopetrosis). OST polypeptides are useful			
CC	CC	as immunogens to raise anti-OST antibodies (II). Parathyroid hormone			
CC	CC	receptor ligands (I) identified by (M1) are useful for treating bone			
CC	CC	disorders. The present sequence represents an OST23 gene fragment			
CC	CC	isolated from rat tibiae, from the present invention.			
XX	SO	Sequence 439 BP; 128 A; 108 C; 99 G; 104 T; 0 other;			
Query Match					
Best Local Similarity 20.7%; Score 354; DB 24; Length 439;					
Matches 404; Conservative 99.8%; Pred. NO. 4e-162; Mismatches 1; Indels 0; Gaps 0;					
OY	1250	TGTACAAAAGAGCAACCAAAAAGACCTTAAACCTAGGCTATATCTGGCAAACTGGCC	1309		
Db	1	TGTACAAAAGAGCAACCAAAAAGACCTTAAACCTAGGCTATATCTGGCAAACTGGCC	60		
OY	1310	ACCGTGTGGGATTAAGTCAATGTTAGGACCAAGACAGACAGATTTGCCGTAACCTTCCAAT	1369		
Db	61	ACCGTGTGGGATTAAGTCAATGTTAGGACCAAGACAGACAGATTTGCCGTAACCTTCCAAT	120		
OY	1370	TCCCTTCTTGACTTCTGTATGCTGTGCCCAAGATGATGAATGAATCTGTAAGTGTAC	1429		
Db	121	TCCCTTCTTGACTTCTGTATGCTGTGCCCAAGATGATGAATGAATCTGTAAGTGTAC	180		
OY	1430	CTTCCCTGACCTGAGAACACCCCTGCTGCGGGAAGTATTCAGGGCAGAATTCTCTGT	1489		
Db	181	CTTCCCTGACCTGAGAACACCCCTGCTGCGGGAAGTATTCAGGGCAGAATTCTCTGT	240		
OY	1490	GAACATGAAGATGAATCACACTGTCTTAAAGAAATTCCTGAAGTCCAGAACTTGAG	1549		
Db	241	GAACATGAAGATGAATCACACTGTCTTAAAGAAATTCCTGAAGTCCAGAACTTGAG	300		
OY	1550	CTTGTATTTTCAAGGAATGCACATCTTTAAGCACTCGCAAAACAGGAAGGCTCCACACC	1609		
Db	301	CTTGTATTTTCAAGGAATGCACATCTTTAAGCACTCGCAAAACAGGAAGGCTCCACACC	360		
OY	1610	TCTGGCAGGCGAGGGCCTTCTCTTACAGATGAGAAAGACAAAGG 1654			
Db	361	TCTGGCAGGCGAGGGCCTTCTCTTACAGATGAGAAAGACAAAGG 405			

XX Rat HICP TSPI domain encoding cDNA.
DE
XX
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder;
KW thrombospondin 1 domain; TSPI.
XX
OS Rattus sp.
XX
PN WO9947556-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Castellot JJ;
XX
DR WPI: 1999-562060/47.
DR P-PSDB; AAY27439.
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications
XX
PS Disclosure; Page 105; 108pp; English.
XX
CC The invention provides a rat heparin-induced CCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents a nucleotide sequence encoding the
CC thrombospondin 1 (TSPI) domain of the HICP polypeptide.
XX
SQ Sequence 174 BP; 36 A; 62 C; 45 G; 31 T; 0 other;
XX
Query Match 10.2%; Score 174; DB 20; Length 174;
Best Local Similarity 100.0%; Pred. No. 2.2e-74;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 825 CCTGTCCAAATTGGAGACAGAGCCCTGGGGCCCTGCTCAACACCTGTGGGCTGGGCATA 884
DB 1 CCTGTCCAAATTGGAGACAGAGCCCTGGGGCCCTGCTCAACACCTGTGGGCTGGGCATA 60
OY 885 GCCACCCGAGTGTCCAAACAGAACCGATTCTGCCAACTGGAGATCCAAAGCCGCTGTGT 944
DB 61 GCCACCCGAGTGTCCAAACAGAACCGATTCTGCCAACTGGAGATCCAAAGCCGCTGTGT 120
OY 945 CTGCCACAGACCTGCTGGCAGCCAGAGCCACAGCTCATGGACAGTCTTTC 998
DB 121 CTGCCACAGACCTGCTGGCAGCCAGAGCCACAGCTCATGGACAGTCTTTC 174
RESULT 8
AAAX76489/c
ID AAX76489 standard; DNA; 753 BP.
XX
AC AAX76489;
XX
DT 06-AUG-1999 (first entry)
XX
DE Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;

KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Mus sp.
XX
PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI: 1999-337420/28.
XX
PT New isolated wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Disclosure; Page 179-180; 284pp; English.
XX
CC The present invention describes wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 753 BP; 159 A; 229 C; 238 G; 127 T; 0 other;
XX
Query Match 5.3%; Score 90; DB 20; Length 753;
Best Local Similarity 100.0%; Pred. No. 1.9e-33;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 410 AGTGTGTACAGGAGGCTGGGGAGTCTGCGACCACTGCATGTCTGGACCCAGCCA 469
DB 592 AGTGTGTACAGGAGGCTGGGGAGTCTGCGACCACTGCATGTCTGGACCCAGCCA 533
OY 470 GGGCTGTGTTTTCAGCCTGGGGCAGGCC 499
DB 532 GGGCTGTGTTTTCAGCCTGGGGCAGGCC 503
RESULT 9
AAAX76488
ID AAX76488 standard; DNA; 1734 BP.
XX
AC AAX76488;
XX
DT 06-AUG-1999 (first entry)
XX
DE Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.

XX Wnt-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Mus sp.
XX
PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WL;
XX
DR WPI; 1999-337420/28.
DR P-PSDB; AAY17651.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Example 2; Page 178-179; 284pp; English.
XX
CC The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoellic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;
XX
Query Match 5.3%; Score 90; DB 20; Length 1734;
Best Local Similarity 100.0%; Pred. No. 1.9e-33;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCTGCATGTCGCGACCCAGCCA 469
|||||
DB 418 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCTGCATGTCGCGACCCAGCCA 477
|||||
QY 470 GGGCTGCTTGTGTCAGCCTGGGGCAGGCC 499
|||||
DB 478 GGGCTGCTTGTGTCAGCCTGGGGCAGGCC 507
|||||

RESULT 10
ABN30189
ID ABN30189 standard; DNA; 65 BP.
XX

AC ABN30189;
XX
DT 15-JUL-2002 (first entry)
XX
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:2937.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Rattus norvegicus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Falgler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
PS Example 1; SEQ ID 2937; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 21 A; 20 C; 13 G; 11 T; 0 other;
XX
Query Match 3.2%; Score 54; DB 24; Length 65;
Best Local Similarity 100.0%; Pred. No. 6.7e-16;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1559 TTCAGAAATGCACATCTCTTAAGCACTCGCAAAACAGAGAGGCTCCACACCTCT 1612
|||||
DB 1 TTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGAGAGGCTCCACACCTCT 54
|||||

RESULT 11
ABA72245
ID ABA72245 standard; DNA; 199 BP.
XX

AC ABA72245;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #20550.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver.
XX
PS Claim 4; SEQ ID NO 20550; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 199 BP; 32 A; 65 C; 73 G; 29 T; 0 other;

Query Match 1.9%; Score 32; DB 22; Length 199;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGGATGTGGCGCTGCCAGCTGGGACTGCC 685
DB 129 GAGGATGTGGCGCTGCCAGCTGGGACTGCC 160

RESULT 12
ABA38112
ID ABA38112 standard; DNA; 199 BP.
XX
AC ABA38112;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #16578 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.

XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 4; SEQ ID NO 16578; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 199 BP; 32 A; 65 C; 73 G; 29 T; 0 other;

Query Match 1.9%; Score 32; DB 22; Length 199;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGGATGTGGCGCTGCCAGCTGGGACTGCC 685
DB 129 GAGGATGTGGCGCTGCCAGCTGGGACTGCC 160

RESULT 13
AAK20667
ID AAK20667 standard; DNA; 199 BP.
XX
AC AAK20667;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 20658.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.

```
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
XX Example 4; SEQ ID NO: 20658; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 199 BP; 32 A; 65 C; 73 G; 29 T; 0 other;

Query Match
Best Local Similarity 1.9%; Score 32; DB 22; Length 199;
Matches 32; Conservative 100.0%; Pred. No. 3.6e-05;
Indels 0; Gaps 0;

QY 654 GAGGATGTGCGGCTGCCAGCTGGGACTGCC 685
Db 129 GAGGATGTGCGGCTGCCAGCTGGGACTGCC 160

RESULT 14
AAK46811
ID AAK46811 standard; DNA; 199 BP.
XX
AC AAK46811;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 21368..
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
```

```
PS Example 4; SEQ ID NO: 21368; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 199 BP; 32 A; 65 C; 73 G; 29 T; 0 other;

Query Match
Best Local Similarity 1.9%; Score 32; DB 22; Length 199;
Matches 32; Conservative 100.0%; Pred. No. 3.6e-05;
Indels 0; Gaps 0;

QY 654 GAGGATGTGCGGCTGCCAGCTGGGACTGCC 685
Db 129 GAGGATGTGCGGCTGCCAGCTGGGACTGCC 160

RESULT 15
AAI25749
ID AAI25749 standard; DNA; 199 BP.
XX
AC AAI25749;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #15682 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID NO 15682; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 199 BP; 32 A; 65 C; 73 G; 29 T; 0 other;
```

Query Match 1.98; Score 32; DB 22; Length 199;
 Best Local Similarity 100.0%; Pred. No. 3.6e-05;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 654 GAGGATGTGGGCTGCCAGCTGGGACTGCCC 685
 ||||||||||||||||||||||||||||||||
 Db 129 GAGGATGTGGGCTGCCAGCTGGGACTGCCC 160

Search completed: July 28, 2003, 15:54:55
 Job time : 452.586 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:50:06 ; Search time 93.3356 Seconds
(without alignments)
5612.050 Million cell updates/sec

Title: US-10-010-408-1
Perfect score: 1708
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GCCTAGATAAACCACCCAAA 1708

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	5.3	1734	4	US-09-182-145-17 Sequence 17, Appl
2	90	5.3	1734	4	US-09-182-145-18 Sequence 18, Appl
3	32	1.9	738	4	US-09-182-145-38 Sequence 38, Appl
4	32	1.9	841	4	US-09-182-145-39 Sequence 39, Appl
5	32	1.9	1293	4	US-09-182-145-13 Sequence 13, Appl
6	32	1.9	1293	4	US-09-182-145-14 Sequence 14, Appl
7	27	1.6	51	4	US-09-182-145-117 Sequence 117, App
8	19	1.1	616	4	US-09-385-982-220 Sequence 220, App
9	19	1.1	1196	4	US-09-149-476-225 Sequence 225, App
10	19	1.1	1220	4	US-09-149-476-57 Sequence 57, Appl
11	19	1.1	1514	2	US-09-213-768-1 Sequence 1, Appl
12	19	1.1	1539	4	US-09-668-680-13 Sequence 13, Appl
13	18	1.1	20	2	US-09-213-768-2 Sequence 2, Appl
14	18	1.1	2104	4	US-09-313-930-1 Sequence 1, Appl
15	18	1.1	2949	4	US-08-259-451-10 Sequence 10, Appl
16	18	1.1	3727	1	US-08-249-380-1 Sequence 1, Appl
17	18	1.1	8957	4	US-08-259-451-1 Sequence 1, Appl
18	17	1.0	24	4	US-09-182-145-110 Sequence 110, App
19	17	1.0	44	4	US-09-182-145-152 Sequence 152, App
20	17	1.0	464	2	US-08-691-814B-117 Sequence 117, App
21	17	1.0	480	3	US-09-188-930-206 Sequence 206, App
22	17	1.0	482	2	US-08-691-814B-120 Sequence 120, App
23	17	1.0	614	4	US-08-998-416-151 Sequence 151, App
24	17	1.0	661	3	US-08-578-634C-3 Sequence 3, Appl
25	17	1.0	661	3	US-08-578-634C-3 Sequence 3, Appl
26	17	1.0	742	4	US-09-430-010-3 Sequence 3, Appl
27	17	1.0	896	3	US-09-188-930-36 Sequence 36, Appl

28	17	1.0	933	3	US-08-987-743-1 Sequence 1, Appl
29	17	1.0	1308	3	US-08-987-743-5 Sequence 5, Appl
30	17	1.0	1491	4	US-09-662-249A-3 Sequence 3, Appl
31	17	1.0	1743	3	US-08-665-259-20 Sequence 20, Appl
32	17	1.0	1743	3	US-08-762-500-20 Sequence 20, Appl
33	17	1.0	1821	4	US-09-149-476-90 Sequence 90, Appl
34	17	1.0	1974	3	US-08-762-500-78 Sequence 78, Appl
35	17	1.0	2329	1	US-08-455-559-9 Sequence 9, Appl
36	17	1.0	2329	4	US-09-145-060-9 Sequence 9, Appl
37	17	1.0	2329	5	PCT-US94-00657-9 Sequence 9, Appl
38	17	1.0	2517	3	US-08-733-360A-4 Sequence 4, Appl
39	17	1.0	2517	3	US-08-916-935-4 Sequence 4, Appl
40	17	1.0	2517	4	US-08-987-743-16 Sequence 16, Appl
41	17	1.0	3280	4	US-09-484-970B-43 Sequence 43, Appl
42	17	1.0	3865	4	US-09-149-476-296 Sequence 296, App
43	17	1.0	6803	3	US-08-665-259-19 Sequence 19, Appl
44	17	1.0	6803	3	US-08-762-500-19 Sequence 19, Appl
45	17	1.0	7874	4	US-09-780-175-96 Sequence 96, Appl

ALIGNMENTS

RESULT 1
US-09-182-145-17
; Sequence 17, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-182-145-17

Query Match 5.3%; Score 90; DB 4; Length 1734;
Best local Similarity 100.0%; Pred. No. 1.9e-34;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTGACGAGGCTGGGGAGTCTGCGACCACTGATGTCTGCGACCCAGCCA 469
DB 418 AGTGTGTGACGAGGCTGGGGAGTCTGCGACCACTGATGTCTGCGACCCAGCCA 477
QY 470 GGGCTGTGTTGTCTGAGCCTGGGGCAGGCC 499
DB 478 GGGCTGTGTTGTCTGAGCCTGGGGCAGGCC 507

RESULT 2
US-09-182-145-18/c
; Sequence 18, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 18
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-182-145-18

Query Match 5.3%; Score 90; DB 4; Length 1734;
Best Local Similarity 100.0%; Pred. No. 1.9e-34;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTCACGAGGCTGGGGAGTCTCGGACGACGTCATGTCGAGCCAGCCCA 469
DB 1317 AGTGTGTCACGAGGCTGGGGAGTCTCGGACGACGTCATGTCGAGCCAGCCCA 1258

QY 470 GGGCTGTGTTGTTCAGCCTGGGGCAGGCC 499
DB 1257 GGGCTGTGTTGTTCAGCCTGGGGCAGGCC 1228

RESULT 3
US-09-182-145-38
; Sequence 38, Application US/09182145B
; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 38
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-38

Query Match 1.9%; Score 32; DB 4; Length 738;

Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGTGCTGCATGGCTGTGGCTGCTG 406
DB 115 GTACCCCTGTGCTGCATGGCTGTGGCTGCTG 146

RESULT 4
US-09-182-145-39

; Sequence 39, Application US/09182145B
; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 39
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Artificial sequence
; NAME/KEY: misc_feature
; LOCATION: 1-841
; OTHER INFORMATION: Sequence is synthesized.
; Patent No. 6387657
US-09-182-145-39

Query Match 1.9%; Score 32; DB 4; Length 841;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGATGTGGCGCTGCCAGCTGGAGCTGCC 685
DB 417 GAGATGTGGCGCTGCCAGCTGGAGCTGCC 448

RESULT 5
US-09-182-145-13

; Sequence 13, Application US/09182145B
; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29

```
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 13
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-13

Query Match          1.9%; Score 32; DB 4; Length 1293;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
Db 148 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 179

RESULT 6
US-09-182-145-14/c
; Sequence 14, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 14
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-14

Query Match          1.9%; Score 32; DB 4; Length 1293;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
Db 1146 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 1115

RESULT 7
US-09-182-145-117
; Sequence 117, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
```

```
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 117
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1-51
; OTHER INFORMATION: Sequence is synthesized.
; Patent No. 6387657
US-09-182-145-117

Query Match          1.6%; Score 27; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 CCTGCTGCTGATGGCTGTGGCTGCTG 406
Db 1 CCTGCTGCTGATGGCTGTGGCTGCTG 27

RESULT 8
US-09-385-982-220/c
; Sequence 220, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 220
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(616)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-220

Query Match          1.1%; Score 19; DB 4; Length 616;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 TGATGACGGTGGCTTCACC 635
Db 127 TGATGACGGTGGCTTCACC 109
```



```

RESULT 9
US-09-149-476-225/C
; Sequence 225, Application US/09149476
; Patent NO. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; EARLIER FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11

```

```

; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595

```

```
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02
```

```
Query Match          1.1%; Score 19; DB 4; Length 1196;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      617 TGATGACGGTGGCTTCACC 635
Db      134 TGATGACGGTGGCTTCACC 116
```

```
RESULT 10
US-09-149-476-57/c
; Sequence 57, Application US/09149476
; Patent No. 6420326
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

```
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; EARLIER FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
```

EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 1.1%; Score 19; DB 4; Length 1220;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 TGATGACGGTGCCTCACC 635
|||||
Db 128 TGATGACGGTGCCTCACC 110

RESULT 11
US-09-213-768-1/c
Sequence 1, Application US/09213768
Patent No. 5985664
GENERAL INFORMATION:
APPLICANT: Brenda F. Baker
APPLICANT: lex M. Cowsett
TITLE OF INVENTION: ANTISENSE MODULATION OF SENTRIN EXPRESSION
FILE REFERENCE: RTS-0026
CURRENT APPLICATION NUMBER: US/09/213,768
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 47
SEQ ID NO 1

; LENGTH: 1514
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(441)
US-09-213-768-1

Query Match 1.1%; Score 19; DB 2; Length 1514;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 TGATGACGGTGGCTTCACC 635
|||||
Db 136 TGATGACGGTGGCTTCACC 118

RESULT 12

US-09-668-680-13
; Sequence 13, Application US/09668680
; Patent No. 6436703
; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Xu, Chongjun
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6436703el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 790CIP2A
; CURRENT APPLICATION NUMBER: US/09/668,680
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 13
; LENGTH: 1539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (130)..(1539)
US-09-668-680-13

Query Match 1.1%; Score 19; DB 4; Length 1539;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 CTTCTCTGCTCTCTCA 305
|||||
Db 774 CTTCTCTGCTCTCTCA 792

RESULT 13

US-09-213-768-2/c
; Sequence 2, Application US/09213768
; Patent No. 5985664
; GENERAL INFORMATION:

; APPLICANT: Brenda F. Baker
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF SENTRIN EXPRESSION
; FILE REFERENCE: RTS-0026
; CURRENT APPLICATION NUMBER: US/09/213,768
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 2

; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR Primer
US-09-213-768-2

Query Match 1.1%; Score 18; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 TGATGACGGTGGCTTCAC 634
|||||
Db 18 TGATGACGGTGGCTTCAC 1

RESULT 14

US-09-313-930-1/c
; Sequence 1, Application US/09313930
; Patent No. 6235723
; GENERAL INFORMATION:

; APPLICANT: Dean, Nicholas M.
; TITLE OF INVENTION: Antisense Oligonucleotide Modulation of Human Protein
; TITLE OF INVENTION: Kinase C-delta Expression
; FILE REFERENCE: ISPH-0357
; CURRENT APPLICATION NUMBER: US/09/313,930
; CURRENT FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2104
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)..(2089)
; PUBLICATION INFORMATION:

; AUTHORS: Aris, J. P.
; AUTHORS: Basta, P. V.
; AUTHORS: Holmes, W. D.
; AUTHORS: Ballas, L. M.
; AUTHORS: Moomaw, C.
; AUTHORS: Rankl, N. B.
; AUTHORS: Blobel, G.
; AUTHORS: Loomis, C. R.
; AUTHORS: Burns, D. J.
; TITLE: Molecular and biochemical characterization of a
; TITLE: recombinant human PKC-delta family member
; JOURNAL: Biochim. Biophys. Acta
; VOLUME: 1174
; ISSUE: 2
; PAGES: 171-181
; DATE: 1993-08-19
; DATABASE ACCESSION NUMBER: L07860
; DATABASE ENTRY DATE: 1993-11-02
US-09-313-930-1

Query Match

Query Match 1.1%; Score 18; DB 4; Length 2104;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1658 GCAGAGTACTCTCTCTG 1675
|||||
Db 1185 GCAGAGTACTCTCTCTG 1168

RESULT 15

US-08-259-451-10
; Sequence 10, Application US/08259451
; Patent No. 6406841
; GENERAL INFORMATION:

; APPLICANT: Lee, Helen H.
; APPLICANT: Swanson, Priscilla A.

```

; APPLICANT: Idler, Kenneth B.
; APPLICANT: Rosenblatt, Joseph D.
; APPLICANT: Chen, Irvin S. Y.
; APPLICANT: Golde, David W.
; APPLICANT: Robertson, Eugene F.
; APPLICANT: Stephens, John E.
; APPLICANT: Chan, Emerson W.
; APPLICANT: Buytendorp, Mark H.
; APPLICANT: Johnson, Joan E.
; APPLICANT: Motley, Cheryl T.
; APPLICANT: Peterson, Bryan
; APPLICANT: Edwards, Michelle
; APPLICANT: Guidinger, Peggy
; APPLICANT: Tate, Cynthia
; TITLE OF INVENTION: HTLV-III-RNA Compositions
; TITLE OF INVENTION: and Assays for Detecting HTLV Infection
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: One Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; MEDIUM TYPE: storage
; COMPUTER: IBM
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,451
; FILING DATE: 20-JUN-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,415
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Daniel W. Collins
; REGISTRATION NUMBER: 31,912
; REFERENCE/DOCKET NUMBER: 5381.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 937-6365
; TELEFAX: (708) 938-2623
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2949 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
; US-08-259-451-10

```

```

Query Match 1.18; Score 18; DB 4; length 2949;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 193 GGCTTCACGGGCTCACC 210
   ||||||||||||||||
Db 2071 GGCTTCACGGGCTCACC 2088

```

Search completed: July 28, 2003, 15:58:52
 Job time : 94.3356 secs

LENGTH: 1708 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 249..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-010-408-1

Query Match 100.0%; Score 1708; DB 15; Length 1708;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGCTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCGCTTGGCAAGGCTGCAGCC 60
Db 1 GACGCTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCGCTTGGCAAGGCTGCAGCC 60
QY 61 GCTGGGCACTGGCTTGGAAATGAGGTCTTTAATTAAGTGGGAACTGAGAGCTAAGAGGCTC 120
Db 61 GCTGGGCACTGGCTTGGAAATGAGGTCTTTAATTAAGTGGGAACTGAGAGCTAAGAGGCTC 120
QY 121 CTGTCAAGCTTGTCTAAAGTCTTAGACATTTGTGGCTTGGGCTTCAACACTGTGAGA 180
Db 121 CTGTCAAGCTTGTCTAAAGTCTTAGACATTTGTGGCTTGGGCTTCAACACTGTGAGA 180
QY 181 CACCTTCGTGTGGCTCCACGGCTCACCTTCAGGTTTGAAGCTGGCTCCACAGGAGAC 240
Db 181 CACCTTCGTGTGGCTCCACGGCTCACCTTCAGGTTTGAAGCTGGCTCCACAGGAGAC 240
QY 241 ACGGTGACATGAGGGGACAGCCCACTGATCCATCTTCTGGCACTTCCCTCTGCTTC 300
Db 241 ACGGTGACATGAGGGGACAGCCCACTGATCCATCTTCTGGCACTTCCCTCTGCTTC 300
QY 301 TCTCAATGGTGTGTGCCCAAGCTGTGGCGAGACCCCTGTAGCTGTCTGTGGACACACCC 360
Db 301 TCTCAATGGTGTGTGCCCAAGCTGTGGCGAGACCCCTGTAGCTGTCTGTGGACACACCC 360
QY 361 AGTCCCACAGGGGGTACCCCTGTGTGTGATGGCTGTGGCTGTGTAAGTGTGTGCAC 420
Db 361 AGTCCCACAGGGGGTACCCCTGTGTGTGATGGCTGTGGCTGTGTAAGTGTGTGCAC 420
QY 421 GGAGGCTGGGGAGTCTCGACACACCTGCAATGTCTGCAACCCAGCCAGGGGCTGTT 480
Db 421 GGAGGCTGGGGAGTCTCGACACACCTGCAATGTCTGCAACCCAGCCAGGGGCTGTT 480
QY 481 GTCAAGCTGGGGAGGCGCTGGCGCATGGGGCTGTGTCTCTTGAATGAGATGACG 540
Db 481 GTCAAGCTGGGGAGGCGCTGGCGCATGGGGCTGTGTCTCTTGAATGAGATGACG 540
QY 541 GTAGCTGTGAGTGAATGGCGCAGGTACTTGATGAGAGACCTTTAAACCAATTGCA 600
Db 541 GTAGCTGTGAGTGAATGGCGCAGGTACTTGATGAGAGACCTTTAAACCAATTGCA 600
QY 601 GGGTCTGTGGCGCTGTGATGACGGTGGCTTCACTGGCGGCTGTGAGTGAAGATG 660
Db 601 GGGTCTGTGGCGCTGTGATGACGGTGGCTTCACTGGCGGCTGTGAGTGAAGATG 660
QY 661 TGGGGTGGCCAGCTGGGACTGCCACGCGCCCAAGAGATACAGGTGCCAGGAAGTGCT 720
Db 661 TGGGGTGGCCAGCTGGGACTGCCACGCGCCCAAGAGATACAGGTGCCAGGAAGTGCT 720
QY 721 GCCCGAGTGGGTATGTGACAGGAGTGAACCGGCGATCCAGCGCTCCACGGCGCAAG 780
Db 721 GCCCGAGTGGGTATGTGACAGGAGTGAACCGGCGATCCAGCGCTCCACGGCGCAAG 780
QY 781 GACACCACTTCTGCCCTTGTCACTCTCTCTGTCTGATGCTCTTGTCCAAATTGGA 840
Db 781 GACACCACTTCTGCCCTTGTCACTCTCTCTGTCTGATGCTCTTGTCCAAATTGGA 840
QY 841 GCACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGATAGCCACCGAGTGTCCA 900
Db 841 GCACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGATAGCCACCGAGTGTCCA 900

Db 841 GCACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGATAGCCACCGAGTGTCCA 900
QY 901 ACCAGAACCGATTTCTGCCAACTGGAGATCCAAACGCGGCTGTGTGCCCCAGACCTGCC 960
Db 901 ACCAGAACCGATTTCTGCCAACTGGAGATCCAAACGCGGCTGTGTGCCCCAGACCTGCC 960
QY 961 TGGCAGCCAGGACCCACAGCTCATGGAACAGTCTTTCTAAGGCCAACTGGGATGCGGA 1020
Db 961 TGGCAGCCAGGACCCACAGCTCATGGAACAGTCTTTCTAAGGCCAACTGGGATGCGGA 1020
QY 1021 TACAGGGCTGCCATCTCTCAGCAAAATGACCTTAGACACAGGCGCTGCACTGTGTAGAT 1080
Db 1021 TACAGGGCTGCCATCTCTCAGCAAAATGACCTTAGACACAGGCGCTGCACTGTGTAGAT 1080
QY 1081 GCTCTTCTCATGCTCTTGGCTGCAATTAAGTGTCTGCTTGAATCACTGTGTAGAGCC 1140
Db 1081 GCTCTTCTCATGCTCTTGGCTGCAATTAAGTGTCTGCTTGAATCACTGTGTAGAGCC 1140
QY 1141 ACTGAGCGATCCCTGCTCTGTGAGTGAAGCGGAGAGAGGTGACACAGCTCCAGTCTCTG 1200
Db 1141 ACTGAGCGATCCCTGCTCTGTGAGTGAAGCGGAGAGAGGTGACACAGCTCCAGTCTCTG 1200
QY 1201 GTTCAGCCTGGAAATCTGGGTTCTCTGCTCATTTCTCTCAAAACATCCCTGTACAAAAG 1260
Db 1201 GTTCAGCCTGGAAATCTGGGTTCTCTGCTCATTTCTCTCAAAACATCCCTGTACAAAAG 1260
QY 1261 GACAACCAAAAAGACCTTTAAACCTAGGCTATACCTGGGCAAACTGGCCACCGTGTGGG 1320
Db 1261 GACAACCAAAAAGACCTTTAAACCTAGGCTATACCTGGGCAAACTGGCCACCGTGTGGG 1320
QY 1321 GATAAGGTCAATGTTAGAGACAGACAGATTTGCTGAACCTTCCCTCTTGG 1380
Db 1321 GATAAGGTCAATGTTAGAGACAGACAGATTTGCTGAACCTTCCCTCTTGG 1380
QY 1381 GATTAAGGTCAATGTTAGAGACAGACAGATTTGCTGAACCTTCCCTCTTGG 1440
Db 1381 GATTAAGGTCAATGTTAGAGACAGACAGATTTGCTGAACCTTCCCTCTTGG 1440
QY 1441 TGAGAACACCTGCTGCTGGGAGTATTCAGGGGAGAAATTTCTGTGAACATGAAGA 1500
Db 1441 TGAGAACACCTGCTGCTGGGAGTATTCAGGGGAGAAATTTCTGTGAACATGAAGA 1500
QY 1501 GATGAATCACTGTCTTAAAGCAATTCCTGAAGTCCAGGAACCTTGTATTTT 1560
Db 1501 GATGAATCACTGTCTTAAAGCAATTCCTGAAGTCCAGGAACCTTGTATTTT 1560
QY 1561 CAGGAATGACATCTCTTAAAGCAATTCCTGAAGTCCAGGAACCTTGTATTTT 1620
Db 1561 CAGGAATGACATCTCTTAAAGCAATTCCTGAAGTCCAGGAACCTTGTATTTT 1620
QY 1621 AGGCGCTTCTCTTCAAGCATGAGAAAGCAAGGAGAGAGAGTCTCTCTGAGAGA 1680
Db 1621 AGGCGCTTCTCTTCAAGCATGAGAAAGCAAGGAGAGAGAGTCTCTCTGAGAGA 1680
QY 1681 CTAGTCTAGCCTAGAATAAACACCCAAA 1708
Db 1681 CTAGTCTAGCCTAGAATAAACACCCAAA 1708

RESULT 2
US-10-010-408-3
; Sequence 3, Application US/10010408
; Publication No. US20020165185A1
; GENERAL INFORMATION:
; APPLICANT: John J. Castellot, Jr.
; TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules
; and Uses Therefor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts

COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 753 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..750
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-010-408-3

Query Match 44.1%; Score 753; DB 15; Length 753;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ATGAGGGGACCCCACTGATCCATCTTCGACCCTTCCTGCTCTCTCAATG 308
1 ATGAGGGGACCCCACTGATCCATCTTCGACCCTTCCTGCTCTCTCAATG 60
QY 309 GTGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGACACACCCAGTGCCA 368
61 GTGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGACACACCCAGTGCCA 120
QY 369 CAGGGGTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 428
121 CAGGGGTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 429 GGGGAGTCTTCGACACCACTGCATGTCTGACACCCAGCCAGGGGCTTGTGACGCT 488
181 GGGGAGTCTTCGACACCACTGCATGTCTGACACCCAGCCAGGGGCTTGTGACGCT 240
QY 489 GGGGAGGCTTCGACACCACTGCATGTCTGACACCCAGCCAGGGGCTTGTGACGCTGT 548
241 GGGGAGGCTTCGACACCACTGCATGTCTGACACCCAGCCAGGGGCTTGTGACGCTGT 300
QY 549 GAGGTGAATGCCCGCAGAGTACCTGATGAGAGACCTTTAAACCAATTGACAGGCTCCTG 608
301 GAGGTGAATGCCCGCAGAGTACCTGATGAGAGACCTTTAAACCAATTGACAGGCTCCTG 360
QY 609 TGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCCTGTGCAGTGAAGATGTGGGCTG 668
361 TGCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCCTGTGCAGTGAAGATGTGGGCTG 420
QY 669 CCCAGCTGGAGTCCCCACGCCCCAAGAGATAACAGGTGCCAGGAAGTGTGCCCCGAG 728
421 CCCAGCTGGAGTCCCCACGCCCCAAGAGATAACAGGTGCCAGGAAGTGTGCCCCGAG 480
QY 729 TGGGTATGTGACGAGGAGTGAACACCGGGGATCCAGCGCTCCACGGGCGCAAGGACACCAA 788

|||||
DB 481 TGGGTATGTGACGAGGAGTGAACACCGGGGATCCAGCGCTCCACGCGCAAGGACACCAA 540
QY 789 CTTTCTGCGCTGTGACTCCTGCTCTGCTGATGCTCCTTGTCCAATTGAGACACAGCC 848
DB 541 CTTTCTGCGCTGTGACTCCTGCTCTGCTGATGCTCCTTGTCCAATTGAGACACAGCC 600
QY 849 TGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACACAGAAC 908
DB 601 TGGGGCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAGTGTCCAACACAGAAC 660
QY 909 CGATTCTGCCAAGTGAAGATCCACGCGGCTGTGTCTGCGCCAGACCCCTGCGACAGCC 968
DB 661 CGATTCTGCCAAGTGAAGATCCACGCGGCTGTGTCTGCGCCAGACCCCTGCGACAGCC 720
QY 969 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 1001
DB 721 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 753

RESULT 3
US-10-010-408-12
Sequence 12, Application US/10010408
Publication No. US20020165185A1
GENERAL INFORMATION:
APPLICANT: John J. Castellot, Jr.
TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules
and Uses Therefor
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)742-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 681 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..681
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-010-408-12
Query Match 39.9%; Score 681; DB 15; Length 681;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 318 CAGCTGTGCCGACACCCCTGTACCTGTCTCTGGACACCAACCCAGTGGCCACAGCGGTA 377
Db 1 CAGCTGTGCCGACACCCCTGTACCTGTCTCTGGACACCAACCCAGTGGCCACAGCGGTA 60
QY 378 CCCCTGTGTGATGGCTGTGGCTGTGTAAAGTGTGTGACAGGAGGCTGGGGAGTCC 437
Db 61 CCCCTGTGTGATGGCTGTGGCTGTGTAAAGTGTGTGACAGGAGGCTGGGGAGTCC 120
QY 438 TGGGACCACTGCATGTCTGCGACCCCAAGCCAGAGGCGCTTGTGTACGCTGGGCGAGCC 497
Db 121 TGGGACCACTGCATGTCTGCGACCCCAAGCCAGAGGCGCTTGTGTACGCTGGGCGAGCC 180
QY 498 CCTGGCGGCCATGGGCTGTGTCTCTTGATGAGATGACGGTAGCTGTGAGTGAAT 557
Db 181 CCTGGCGGCCATGGGCTGTGTCTCTTGATGAGATGACGGTAGCTGTGAGTGAAT 240
QY 558 GGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCAATTGACGGTCCCTGTGCGCTGT 617
Db 241 GGGCCGAGGTACCTGGATGGAGAGACCTTTAAACCAATTGACGGTCCCTGTGCGCTGT 300
QY 618 GATGACGGTGTCTACCTGCTGCGCTGTGTGACGTAGATGTGCGGCTGCCAGCTGG 677
Db 301 GATGACGGTGTCTACCTGCTGCGCTGTGTGACGTAGATGTGCGGCTGCCAGCTGG 360
QY 678 GACTGCCACAGCCCCAAGAGATACAGGTGCCAGAAAGTGTGCCCCGAGTGGTATGT 737
Db 361 GACTGCCACAGCCCCAAGAGATACAGGTGCCAGAAAGTGTGCCCCGAGTGGTATGT 420
QY 738 GACCAGGAGTGACACCGCGCATCCAGCGCTCCACGCGCAAGGACACCAACTTCTGCC 797
Db 421 GACCAGGAGTGACACCGCGCATCCAGCGCTCCACGCGCAAGGACACCAACTTCTGCC 480
QY 798 CTGTGCTACTCTGCTCTGTGTGATGCTCTCTGTCCAAATGGAGACACAGCCTGGGCCCC 857
Db 481 CTGTGCTACTCTGCTCTGTGTGATGCTCTCTGTCCAAATGGAGACACAGCCTGGGCCCC 540
QY 858 TGCTCAACCACTGTGGGCTGGGCGATAGCCACCCGAGTGTCCAAACAGAACCGATCTGC 917
Db 541 TGCTCAACCACTGTGGGCTGGGCGATAGCCACCCGAGTGTCCAAACAGAACCGATCTGC 600
QY 918 CAACCTGGAGATCCAAACGCGCTGTGTCTGTGCGCAGACCTGCTGCGACGACGAGCCAC 977
Db 601 CAACCTGGAGATCCAAACGCGCTGTGTCTGTGCGCAGACCTGCTGCGACGAGCCAC 660
QY 978 AGCTCATGGAACAGTGTCTTC 998
Db 661 AGCTCATGGAACAGTGTCTTC 681
```

RESULT 4

```
US-09-956-622A-23
; Sequence 23, Application US/09956622A
; Publication No. US20030091973A1
; GENERAL INFORMATION:
; APPLICANT: Horesovsky, Gregory J
; APPLICANT: No. US20030091973A11 II, L. Staton
; APPLICANT: Raha, Debasish
; TITLE OF INVENTION: Method of Identifying Osteoregenerative Agents Using
; FILE REFERENCE: 21402-445
; CURRENT APPLICATION NUMBER: US/09/956,622A
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/233,579
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-956-622A-23
```

```
Query Match 20.7%; Score 354; DB 12; Length 439;
Best Local Similarity 99.8%; Pred. No. 3e-178;
Matches 404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1250 TGTACAAAAGGACCAACCAAAAAGACCTTTAAACCTAGGCTATACTGGGCAACCTGGCC 1309
Db 1 TGTACAAAAGGACCAACCAAAAAGACCTTTAAACCTAGGCTATACTGGGCAACCTGGCC 60
QY 1310 ACCGTGCTGGGATAGGTCAATGTTAGACACAGACAGATTTGCCGTGAACCTTCCAAT 1369
Db 61 ACCGTGCTGGGATAGGTCAATGTTAGACACAGACAGATTTGCCGTGAACCTTCCAAT 120
QY 1370 TCCCTTCTTGACTTCTGTATGCTTGTCCCAAGATGATGAATGAACCTGTAAGTGTAC 1429
Db 121 TCCCTTCTTGACTTCTGTATGCTTGTCCCAAGATGATGAATGAACCTGTAAGTGTAC 180
QY 1430 CTTCCCTGACCTGAGAACACCCCTGCCCTGCTCGGAGATATTCAGGGCGAGAAATCTCTGT 1489
Db 181 CTTCCCTGACCTGAGAACACCCCTGCCCTGCTCGGAGATATTCAGGGCGAGAAATCTCTGT 240
QY 1490 GAACATGAAGATGATCACAACCTGCTCTTAAGAAATTCCTGAAGTCCAGAACTTGAG 1549
Db 241 GAACATGAAGATGATCACAACCTGCTCTTAAGAAATTCCTGAAGTCCAGAACTTGAG 300
QY 1550 CTTGTATTTTCAAGAAATGACATCTCTTAAGACATCGCAAAACAGGAAGCTCCACACC 1609
Db 301 CTTGTATTTTCAAGAAATGACATCTCTTAAGACATCGCAAAACAGGAAGCTCCACACC 360
QY 1610 TCTGCGAGCGCCAGGCGCTTCTCTTACGATGAGAAAGACAAGGG 1654
Db 361 TCTGCGAGCGCCAGGCGCTTCTCTTACGATGAGAAAGACAAGGG 405
```

RESULT 5

```
US-10-010-408-8
; Sequence 8, Application US/10010408
; Publication No. US20020165185A1
; GENERAL INFORMATION:
; APPLICANT: John J. Castellet, Jr.
; TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CGN-Like Molecules
; and Uses Therefor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/010,408
; FILING DATE: 07-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/044,273
; FILING DATE: March 19, 1998
; APPLICATION NUMBER: <unknown>
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MBI-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 base pairs
```


TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..210
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-010-408-8

Query Match 12.3%; Score 210; DB 15; Length 210;
Best Local Similarity 100.0%; Pred. No. 2.2e-101;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 CAGCTGCCCCGACACCCCTGTACTGTCTTGACACACCCCACTGCCACAGGGGTA 377
DB 1 CAGCTGCCCCGACACCCCTGTACTGTCTTGACACACCCCACTGCCACAGGGGTA 60
QY 378 CCCCTGCTGATGCTGTGGCTGTAAAGTGTGTGACAGGAGCTGGGGAGTCC 437
DB 61 CCCCTGCTGATGCTGTGGCTGTAAAGTGTGTGACAGGAGCTGGGGAGTCC 120
QY 438 TGCAGCACCTGCATGTCTGCGACCCAGCGGCTGTGTCTGAGCCTGGGAGGC 497
DB 121 TGCAGCACCTGCATGTCTGCGACCCAGCGGCTGTGTCTGAGCCTGGGAGGC 180
QY 498 CCTGGCGCCATGGGGCTGTGTCTCTTG 527
DB 181 CCTGGCGCCATGGGGCTGTGTCTCTTG 210

RESULT 6
US-10-010-408-5

Sequence 5, Application US/10010408
Publication No. US20020165185A1

GENERAL INFORMATION:

APPLICANT: John J. Castellot, Jr.
TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules
and Uses Therefor

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 177 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..177
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-010-408-5

Query Match 10.4%; Score 177; DB 15; Length 177;
Best Local Similarity 100.0%; Pred. No. 8.9e-84;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 TGTAGGTGATGGCCGACAGTCTGATGAGAGACCTTTAAACCAATTGCAGGTC 605
DB 1 TGTAGGTGATGGCCGACAGTCTGATGAGAGACCTTTAAACCAATTGCAGGTC 60
QY 606 CTGTGCCCTGTGATGACGGTGGCTTCACCTGCGCTGCCCTGTGTGAGAGATGCGG 665
DB 61 CTGTGCCCTGTGATGACGGTGGCTTCACCTGCGCTGCCCTGTGTGAGAGATGCGG 120
QY 666 CTGCCAGCTGGGACTGCCACGCCCCAAGAAGATACAGTGCAGGAAGTGTGC 722
DB 121 CTGCCAGCTGGGACTGCCACGCCCCAAGAAGATACAGTGCAGGAAGTGTGC 177

RESULT 7

US-10-010-408-10

Sequence 10, Application US/10010408
Publication No. US20020165185A1

GENERAL INFORMATION:

APPLICANT: John J. Castellot, Jr.
TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules
and Uses Therefor

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 174 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..174
SEQUENCE DESCRIPTION: SEQ ID NO: 10:


```

: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 23432
: LENGTH: 199
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL139352.8
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
: OTHER INFORMATION: NT HIT: AF083500.1, EVALUE 1.00e-108
: OTHER INFORMATION: SWISSPROT HIT: O19113, EVALUE 9.00e-19
US-09-864-761-23432

Query Match 1.9%; Score 32; DB 10; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGGATGTGGCGCTGCCACAGCTGGGACTGCCC 685
|||||
DB 129 GAGGATGTGGCGCTGCCACAGCTGGGACTGCCC 160

RESULT 11
US-09-864-761-6698
: Sequence 6698, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecm1ca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6698
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL139352.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
US-09-864-761-6698

Query Match 1.9%; Score 32; DB 10; Length 586;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGGATGTGCGCTGCCAGCTGGGACTGCC 685
|||||
DB 342 GAGGATGTGCGCTGCCAGCTGGGACTGCC 373

RESULT 12
US-10-112-267-38
; Sequence 38, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR APPLICATION NUMBER: 2002-03-27
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29

```

```
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 38
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-112-267-38

Query Match      1.9%; Score 32; DB 15; Length 738;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      375 GTACCCCTGTGCTGATGCTGTGCTGTG 406
Db      115 GTACCCCTGTGCTGATGCTGTGCTGTG 146

RESULT 13
US-10-112-267-39
; Sequence 39, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 39
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1-841
; OTHER INFORMATION: Sequence is synthesized.
; US-10-112-267-39

Query Match      1.9%; Score 32; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      654 GAGGATGTGGCGCTGCCAGCTGGAGTGGCC 685
Db      417 GAGGATGTGGCGCTGCCAGCTGGAGTGGCC 448

RESULT 14
US-10-137-866-319
; Sequence 319, Application US/10137866
; Publication No. US20030129689A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
```

```
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C151
; CURRENT APPLICATION NUMBER: US/10/137,866
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
```

; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323

; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982

Query Match 1.9%; Score 32; DB 14; Length 1266;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGCTGTGGCTGCTG 406
|||
DB 136 GTACCCCTGCTGCTGATGCTGTGGCTGCTG 167

RESULT 15
US-10-146-726-319
; Sequence 319, Application US/10146726
; Publication No. US20030129690A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey


```

: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3330R1C308
: CURRENT APPLICATION NUMBER: US/10/146,726
: CURRENT FILING DATE: 2002-05-15
: Prior Application removed - See file wrapper or pajm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 319
: LENGTH: 1266
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-146-726-319

```

Query Match	1.98;	Score 32;	DB 14;	Length 1266;
Best Local Similarity	100.0%;	Pred. No. 1.9e-06;		
Matches 32;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	375	GTACCCCTGTGCTGCATGCGCTGTGCTGCTG	406	
Db	136	GTACCCCTGTGCTGCATGCGCTGTGCTGCTG	167	

Search completed: July 28, 2003, 15:36:31
Job time : 426.627 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:16:26 ; Search time 3029.91 Seconds
(without alignments)
9129.604 Million cell updates/sec

Title: US-10-010-408-1

Perfect score: 1708

Sequence: 1 GACGCTTCTGATCTCCAGAG.....GCCTAGATAACACCAAA 1708

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	501	29.3	762	BQ195526	BQ195526 UI-R-CNL-
2	90	5.3	940	BQ937887	BQ937887 AGENCOURT
3	67	3.9	537	BQ560868	BQ560868 H4067A01-
4	49	2.9	424	BB849097	BB849097 BB849097
5	35	2.0	315	BB374499	BB374499 BB374499
6	35	2.0	369	BB792544	BB792544 BB792544

7	35	2.0	418	9	AI225477	AI225477 ue88b01.Y
8	35	2.0	488	9	AA754979	AA754979 vu55908.r
9	35	2.0	792	12	BF138093	BF138093 601785003
10	35	2.0	966	12	BF141695	BF141695 601790752
11	33	1.9	426	9	AA717584	AA717584 vt97c08.r
12	32	1.9	380	12	BG900020	BG900020 HOA48-1-G
13	32	1.9	405	12	BG900069	BG900069 HOA51-1-A
14	32	1.9	489	14	BM751866	BM751866 K-EST0028
15	32	1.9	618	12	BG538695	BG538695 602566932
16	32	1.9	620	13	BG928868	BG928868 HNC57-1-D
17	32	1.9	651	13	BI457141	BI457141 603185392
18	32	1.9	749	9	AL555144	AL555144 AL555144
19	32	1.9	750	13	BM043988	BM043988 603620978
20	32	1.9	790	13	BM046275	BM046275 603626068
21	32	1.9	800	13	BI826781	BI826781 603077268
22	32	1.9	886	13	BI822142	BI822142 603039845
23	32	1.9	888	13	BI825652	BI825652 603072631
24	32	1.9	916	13	BI457367	BI457367 603185689
25	32	1.9	928	13	BI161474	BI161474 602864871
26	32	1.9	933	14	BQ278961	BQ278961 AGENCOURT
27	32	1.9	979	14	BQ279131	BQ279131 AGENCOURT
28	32	1.9	1006	14	BM921531	BM921531 AGENCOURT
29	32	1.9	1022	14	BQ952960	BQ952960 AGENCOURT
30	32	1.9	1058	14	BM805088	BM805088 AGENCOURT
31	32	1.9	1073	14	BQ073722	BQ073722 AGENCOURT
32	32	1.9	1166	13	BM543799	BM543799 AGENCOURT
33	32	1.9	1251	14	BQ961357	BQ961357 AGENCOURT
34	30	1.8	190	9	AA647775	AA647775 vp04c09.r
35	23	1.3	436	17	AQ095651	AQ095651 HS_3017_A
36	23	1.3	742	13	BI758148	BI758148 603023866
37	23	1.3	792	13	BI823598	BI823598 603040962
38	22	1.3	307	10	BB220676	BB220676 BB220676
39	21	1.2	495	13	BI204749	BI204749 EST522789
40	21	1.2	529	9	AI897896	AI897896 EST267339
41	21	1.2	561	12	BF051668	BF051668 EST436915
42	21	1.2	563	9	AI897344	AI897344 EST266787
43	21	1.2	594	13	BI682551	BI682551 463666 MA
44	21	1.2	609	9	AI485142	AI485142 EST243446
45	21	1.2	620	10	AW223381	AW223381 EST300192

ALIGNMENTS

RESULT 1
BQ195526/c 762 bp mRNA linear EST 30-APR-2002
LOCUS UI-R-CNL-cmq-k-07-0-UI.s1 UI-R-CNL Rattus norvegicus cDNA clone
DEFINITION UI-R-CNL-cmq-k-07-0-UI 3', mRNA sequence.
ACCESSION BQ195526 GI:20371077
VERSION BQ195526.1
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 762)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msquares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized cervix library cDNA Library Preparation: M.B. Soares Lab. Clone distribution: clones will be available through Research Genetics (www.resgen.com) Seq primer: M13 Forward POLYA=yes

FEATURES
source

Location/Qualifiers
1. 762
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CNI-cmq-k-07-0-UI"
/clone_1lb="UI-R-CNI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CNI library is a subtracted library derived from the following pool of seven normalized rat libraries: normalized rat seminal vesicles, normalized rat penis, normalized rat bladder, normalized rat cervix, normalized rat brown adipose, normalized rat fundus, and normalized rat salivary gland. It was constructed according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the CNI library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plamid DNA template preparation) comprising: a) a pool of about 34,000 clones from the Rat Unigene Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CA0 and CA1 corresponding to plates R-CA0-AWV through R-CA0-AXS, R-CA0-AZX through R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS, R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA, R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA1-BBA through R-CA1-BDA, R-CA1-BHZ through R-CA1-BJF, R-CA1-BJR, R-CA1-BJT through R-CA1-BKB, R-CA1-BKD, R-CA1-BKF, R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLH through R-CA1-BLN, R-CA1-BLS, R-CA1-BLD-V, R-CA1-BNR, and R-CA1-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0, CT0, CU0, CW0, and CX0 and normalized libraries CS0, CT0, CU0, CW0, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVV through R-CW0-BWP, R-CW0-BXN through R-CW0-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BV0p (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BFB, R-BV0-ANK through R-BV0-ANR, R-BV0p-AOI through R-BV0p-AOX, and R-BX0-AQY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer

including CV0, CZ0, DA0, DB0, DC0, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYI, R-CZ0-BZB-C, R-DA0-BYJ through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ through R-DB0-BZA, R-DC0-BZI through R-DC0-BZQ, R-DC0-CAY through R-DC0-CBA, R-DD0-BZR through R-DD0-CAA, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CN0, that makes up the tracer. The corresponding plates are R-CNO-BKW through R-CNO-BLD, R-CNO-BLG, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT, R-CNO-BLW-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CN0 pool corresponding to the following addresses: bkw-a-09-0-UI, bkw-b-09-0-UI, bkw-b-11-0-UI, bkw-h-10-0-UI, bkw-d-01-0-UI, bkw-d-06-0-UI, bkw-g-08-0-UI, bkw-h-12-0-UI, bky-a-05-0-UI, bky-a-06-0-UI, bky-a-11-0-UI, bky-c-09-0-UI, bky-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-e-95-0-UI, bld-1-08-0-UI, bld-f-02-0-UI, blq-h-04-0-UI, blr-a-05-0-UI, blt-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CN0 clone (corresponding to the address bky-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.
TAG_LIB=UI-R-CNI
TAG_TISSUE=cervix
TAG_SEQ=GACCA"

BASE COUNT 177 a 176 c 198 g 211 t
ORIGIN

Query Match 29.3%; Score 501; DB 14; Length 762;
Best Local Similarity 99.5%; Pred. No. 1.6e-238;
Matches 651; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1033	CATCCTCAGCAATGACCCCTAGGACGAGGCCCTGAGCTGCTGTAGATGCTCTTCAT	1092
DB	703	CATCCTCAGCAATGACCCCTAGGACGAGGCCCTGAGCTGCTGTAGATGCTCTTCAT	644
QY	1093	GCTCTGGCTGACGTTAACTGCTCTGCTGTGATTCAGTGTAGAGCCACTGAGCGATCC	1152
DB	643	GCTCTGGCTGACGTTAACTGCTCTGCTGTGATTCAGTGTAGAGCCACTGAGCGATCC	584
QY	1153	CTGCTCTGTGAGGTAGGCGGAGCAGGTGACCACTTCAGTCTCTGTGACCTGGA	1212
DB	583	CTGCTCTGTGAGGTAGGCGGAGCAGGTGACCACTTCAGTCTCTGTGACCTGGA	524
QY	1213	ATTCTGGGTTCTCTGGCTCATTCCTCAAAACATCCCTGTACAAAAAGACCAAAAA	1272
DB	523	ATTCTGGGTTCTCTGGCTCATTCCTCAAAACATCCCTGTACAAAAAGACCAAAAA	464
QY	1273	GACCTTTAAACCTAGGCTATACCTGGGCAAACTGGCCACCGTGTGGGATAGTCAAT	1332
DB	463	GACCTTTAAACCTAGGCTATACCTGGGCAAACTGGCCACCGTGTGGGATAGTCAAT	404
QY	1333	GTTAGGACGACAGCATTTGCCGTAACCTTCCATTTCCCTTCTTGACCTTCTGTATGC	1392
DB	403	GTTAGGACGACAGCATTTGCCGTAACCTTCCATTTCCCTTCTTGACCTTCTGTATGC	344
QY	1393	TGTGCCCAAGATGATGAATGAACCTGTAAGTGTACCTTCCCTGACCTGAGACACCCCT	1452
DB	343	TGTGCCCAAGATGATGAATGAACCTGTAAGTGTACCTTCCCTGACCTGAGACACCCCT	284
QY	1453	GCCTGCTCGGAAGTATTCAGGGGAGAGATTCCTGTGAACATGAAGAGATGAATCACAC	1512
DB	283	GCCTGCTCGGAAGTATTCAGGGGAGAGATTCCTGTGAACATGAAGAGATGAATCACAC	224
QY	1513	TGTCCTTAGAATTCCTGAAGTCCAGGACCTGACCTTGTATTTTCAGGAATGCACA	1572

Db 223 TGTCCTTAAGAAATTCCTCAAGTCCAGGAAGCTGAGCTTTGATTTTCAGGAATGCACA 164

QY 1573 TCCTTAAGCACTCGCAAAACAGAGAAGGCTCCACACCTCTGGCAGGCCAGGGCCTTCTC 1632

Db 163 TCCTTAAGCACTCGCAAAACAGAGAAGGCTCCACACCTCTGGCAGGCCAGGGCCTTCTC 104

QY 1633 TTCAGCATGAGAAAGACAGAGGAGCAGAGTACTCTCCTCTGGAGGAGTACTGTC 1686

Db 103 TTCAGCATGAGAAAGACAGAGGAGCAGAGTACTCTCCTCTGGAGGAGTACTGTC 50

RESULT 2

BQ937887

LOCUS BQ937887 940 bp mRNA linear EST 21-AUG-2002

DEFINITION AGENCOURT 8951807 NCI_CGAP_Co24 Mus musculus CDNA clone

IMAGE:6476852 5', mRNA sequence.

ACCESSION BQ937887

VERSION BQ937887.1 GI:22353365

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 940)

TITLE NIH-MGC http://mgc.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: LLAM14017 row: n column: 21

High quality sequence stop: 543.

FEATURES

source 1..940

location/Qualifiers

1..940

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:6476852"

/clone_lib="NCI_CGAP_Co24"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 169 a 277 c 288 g 200 t 6 others

ORIGIN

Query Match 5.3%; Score 90; DB 14; Length 940;

Best Local Similarity 100.0%; Pred. No. 1.8e-33;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCACTGCATGCTGCGACCCAGCCA 469

Db 444 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCACTGCATGCTGCGACCCAGCCA 503

QY 470 GGGCCTGGTTGTCAGCCTGGGGCAGGCC 499

Db 504 GGGCCTGGTTGTCAGCCTGGGGCAGGCC 533

RESULT 3

BQ560868

LOCUS BQ560868 537 bp mRNA linear EST 20-JUN-2002

DEFINITION H4067A01-5 NIA Mouse 7.4K CDNA Clone Set Mus musculus CDNA clone

H4067A01 5', mRNA sequence.

ACCESSION BQ560868

VERSION BQ560868.1 GI:21461753

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 537)

Vanburen,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G., Martin ,P.R., Stagg,C.A., Bassey,U., Alba,K., Hamatani,T., Kargul,G.J., Luo,A.G. and Ko,M.S.H.

Assembly, verification, and initial annotation of NIA 7.4K mouse CDNA clone set

CDNA clone set

Unpublished (2002)

Contact: Yong Qian

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA

Email: cdna@lsun.grc.nia.nih.gov

This clone set has been freely distributed to the community. Please visit http://lgsun.grc.nia.nih.gov/cdna/NIA_7_4k.html for details.

Plate: H4067 row: A column: 01

Seq primer: -21M13 Reverse

High quality sequence stop: 537

POLYA=No.

FEATURES

source 1..537

location/Qualifiers

1..537

/organism="Mus musculus"

/strain="C57BL/6"

/db_xref="niaEST:H4067A01-5"

/db_xref="taxon:10090"

/clone="H4067A01"

/clone_lib="NIA Mouse 7.4K CDNA Clone Set"

/sex="mixed"

/dev_stage="mixed"

/lab_host="DH10B"

/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 7,407 clones from more than 20 CDNA libraries."

BASE COUNT 87 a 162 c 166 g 122 t

ORIGIN

Query Match 3.9%; Score 67; DB 14; Length 537;

Best Local Similarity 100.0%; Pred. No. 4.9e-22;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTCACGAGGCTGGGGAGTCTCGACCACTGCATGCTGCGACCCAGCCA 469

Db 471 AGTGTGTCACGAGGCTGGGGAGTCTCGACCACTGCATGCTGCGACCCAGCCA 530

QY 470 GGGCCTG 476

Db 531 GGGCCTG 537

RESULT 4

BB849097

LOCUS BB849097 424 bp mRNA linear EST 26-NOV-2001

DEFINITION BB849097 RIKEN full-length enriched, adult inner ear Mus musculus

CDNA clone F930006G02 5', mRNA sequence.

ACCESSION BB849097

VERSION BB849097.1 GI:17090551

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 424)

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Orido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa

,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y., RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)

JOURNAL
COMMENT

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES
source

Location/Qualifiers
1. 424
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930006G02"
/clone_lib="RIKEN full-length enriched, adult inner ear"
/tissue_type="inner ear"
/dev_stage="adult"
/note="pooled tissues : (tissue_type=cerebellum, dev_stage=16 days neonate, sex=mixed), (tissue_type=cerebellum, dev_stage=0 day neonate, sex=mixed), (tissue_type=hippocampus, dev_stage=adult, sex=male), (tissue_type=whole body, dev_stage=9 days embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days embryo, sex=mixed)"

BASE COUNT
ORIGIN
65 a 124 c 132 g 103 t

Query Match 2.9%; Score 49; DB 10; Length 424;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 410 AGTGTGTGACGAGGCTGGGGAGTCTGCGACCACTGTCATGTCGC 458
|||||
Db 373 AGTGTGTGACGAGGCTGGGGAGTCTGCGACCACTGTCATGTCGC 421

RESULT 5
BB374499
LOCUS BB374499 315 bp mRNA linear EST 13-JUL-2000
DEFINITION BB374499 RIKEN full-length enriched, 16 days embryo head Mus
musculus cDNA clone C130075G16 3' similar to AF126063 Mus musculus
connective tissue growth factor-like protein precursor (Ctgfl) mRNA
, mRNA sequence.
ACCESSION BB374499
VERSION BB374499.1 GI:9086993
KEYWORDS EST.
SOURCE house mouse.

ORGANISM
REFERENCE
AUTHORS

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 315)
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arahawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsuina,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES
source

Location/Qualifiers
1. 315
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="C130075G16"
/clone_lib="RIKEN full-length enriched, 16 days embryo head"
/sex="mixed"
/tissue_type="head"
/dev_stage="16 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGAAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGATCTCTGAGTATTAATTAATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda
FLC I"]

BASE COUNT
ORIGIN
87 a 80 c 67 g 81 t

Query Match 2.0%; Score 35; DB 10; Length 315;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1610 TCTGGCAGGCCAGGCTTCTCTCAGCATGAGA 1644
|||||
Db 207 TCTGGCAGGCCAGGCTTCTCTCAGCATGAGA 241

RESULT 6
BB792544 369 bp mRNA linear EST 16-NOV-2001
LOCUS BB792544 RIKEN full-length enriched, kidney CCL-142 RAG CDNA Mus
DEFINITION musculus cDNA clone G430141015 3', mRNA sequence.
ACCESSION BB792544
VERSION BB792544
KEYWORDS BB792544.1 GI:16961776
SOURCE EST.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 369)
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

TITLE Unpublished (2001)
JOURNAL Contact: Yoshihide Hayashizaki
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagl,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES
Source Location/Qualifiers
1..369
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="G430141015"
/clone_lib="RIKEN full-length enriched, kidney CCL-142 RAG
CDNA"
/tissue_type="kidney"
/cell_line="CCL-142 RAG"
/note="pooled cell lines : (cell_line=CRL-1751 WEHI 164),
(cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3),

(cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
(cell_line=RCB-0559 K-1.F1), (cell_line=RCB-1283 B16
melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231
) , (cell_type=Leydig cells, cell_line=CRL-2065 MLTC-1),
(cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE),
(tissue_type=bladder, cell_line=RCB-0544 MBF-2),
(tissue_type=bone marrow, cell_type=stroma cell,
cell_line=RCB-2028 SR-4987), (tissue_type=colon,
cell_line=RCB-0549 Cle-H3), (tissue_type=kidney,
cell_line=CCL-142 RAG), (tissue_type=submandibular gland,
cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C,
cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3
) , (strain=C3H, tissue_type=brain, cell_line=CRL-1443
BC3H1)"

BASE COUNT i02 a 90 c 86 g 91 t

ORIGIN

Query Match 2.0%; Score 35; DB 10; Length 369;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1610 TCTGGCAGGCCAGGCTTCTCTCAGCATGAGA 1644
|||||
Db 262 TCTGGCAGGCCAGGCTTCTCTCAGCATGAGA 296

RESULT 7
AI225477 418 bp mRNA linear EST 29-OCT-1998
LOCUS AI225477
DEFINITION ue8b01.y1 Soares_NMPu Mus musculus cDNA clone IMAGE:1498153 5',
mRNA sequence.
ACCESSION AI225477
VERSION AI225477.1 GI:3808530
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 418)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicinep
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:935757
Seq primer: -40RP from Gibco
High quality sequence stop: 403.

FEATURES
Source Location/Qualifiers
1..418
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1498153"
/clone_lib="Soares_NMPu"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; 1st strand cDNA was prepared from
pregnant mouse uterus, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM9276 row: h column: 13
High quality sequence stop: 613.

-FEATURES
source
1. 966
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4021308"
/clone_lib="NCI_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: lung; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 227 a 298 c 258 g 183 t
ORIGIN

Query Match 2.0%; Score 35; DB 12; Length 966;
Best Local Similarity 100.0%; Pred. No. 5e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 897 TCACACGAGACGATCTGCCACTGGAGATCCA 931
|||||
Db 90 TCACACGAGACGATCTGCCACTGGAGATCCA 124

RESULT 11
AA717584 426 bp mRNA linear EST 29-DEC-1997
LOCUS vt97c08.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
IMAGE:1179086 5', mRNA sequence.
ACCESSION AA717584
VERSION AA717584.1 GI:2729858
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 426)
REFERENCE Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:636934
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 416.

FEATURES
Location/Qualifiers

source
1. 426
/organism="Mus musculus"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1179086"
/clone_lib="Soares_mammary_gland_NbMMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer {5',
TGTTACCAATCTGAGCTGGAGCGCGCGCAATGCTTTTCTTTTCTTTTCTTTT
T 3'}; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

BASE COUNT 87 a 123 c 116 g 100 t
ORIGIN

Query Match 1.9%; Score 33; DB 9; Length 426;
Best Local Similarity 100.0%; Pred. No. 4.3e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1086 TCTCCATGCTCTTGGCTGCAGTTAAGTCTG 1118
|||||
Db 199 TCTCCATGCTCTTGGCTGCAGTTAAGTCTG 231

RESULT 12
BG900020 380 bp mRNA linear EST 06-NOV-2001
LOCUS HOA48-1-G2.R HOA (Human Osteoarthritic Cartilage) Homo sapiens cDNA
DEFINITION , mRNA sequence.
ACCESSION BG900020
VERSION BG900020.1 GI:14310269
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
1 (bases 1 to 380)
Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
Sathe,G., Mul,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
Lark,M.W.
Identification and initial characterization of 5000 expressed
sequenced tags (ESTs) each from adult human normal and
osteoarthritic cartilage cDNA libraries
Osteoarthr. Cartil. 9 (7), 641-653 (2001)
MEDLINE 21482651
JOURNAL Contact: Sanjay Kumar
COMMENT UW2109
GlaxoSmithKline
709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
Tel: 610-270-7245
Fax: 610-270-5598
Email: sanjay_kumar-1@gsk.com
Seq primer: T7.

FEATURES
source
1. 380
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HOA (Human Osteoarthritic Cartilage)"
/tissue_type="cartilage"
/lab_host="E.coli DH10 B"
/note="Vector: pSPORT I; Site_1: SalI; Site_2: NotI;
Directional"

BASE COUNT 58 a 139 c 118 g 65 t
ORIGIN

Query Match 1.9%; Score 32; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
|||||
Db 206 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 237

RESULT 13
BG900069 405 bp mRNA linear EST 06-NOV-2001
LOCUS HOA51-1-A11.R HOA (Human Osteoarthritic Cartilage) Homo sapiens
DEFINITION CDNA, mRNA sequence.
ACCESSION BG900069 GI:14310318
VERSION BG900069.1 GI:14310318
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 405)
Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
Sathe,G., Mul,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
Lark,M.W.
TITLE Identification and initial characterization of 5000 expressed
sequenced tags (ESTs) each from adult human normal and
osteoarthritic cartilage cDNA libraries
Osteoarthr. Cartil. 9 (7), 641-653 (2001)
JOURNAL 21482651
MEDLINE
COMMENT Contact: Sanjay Kumar
UM2109
GlaxoSmithKline
709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
Tel: 610-270-7245
Fax: 610-270-5598
Email: sanjay_kumar-1@sk.com
Seq primer: T7.
FEATURES
source Location/Qualifiers
1..405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HOA (Human Osteoarthritic Cartilage)"
/tissue_type="cartilage"
/lab_host="E.coli DH10 B"
/note="Vector: pSPORT I; Site_1: Salt; Site_2: NotI;
Directional"
BASE COUNT 62 a 140 c 135 g 68 t
ORIGIN
Query Match 1.9%; Score 32; DB 12; Length 405;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
|||||
Db 202 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 233

RESULT 14
BM751866 489 bp mRNA linear EST 04-MAR-2002
LOCUS K-EST0028180 S2SNU668s1 Homo sapiens CDNA clone S2SNU668s1-2-D09
DEFINITION 5', mRNA sequence.
ACCESSION BM751866
VERSION BM751866.1 GI:19081484
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 489)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@email.kribb.re.kr
Plate: 2 row: D column: 09
High quality sequence stop: 489.
FEATURES
source Location/Qualifiers
1..489
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S2SNU668s1-2-D09"
/clone_lib="S2SNU668s1"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-668"
/lab_host="Top10F"
/note="Organ: Stomach; Vector: pcNS; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
After analyzing and sequencing about 2,000 - 3,000
colonies in original cDNA library, the abundant cDNAs were
selected and amplified by PCR reaction using vector region
primer including T7 promoter as 5' primer and N(dt)14 as
3' primer. The PCR products were used as template for
synthesis of biotinylated single stranded RNA by in vitro
transcription reaction. The synthesized RNA probes were
hybridized with antisense single stranded cDNAs prepared
from original library and incubated with avidin-gel.
After removing DNA-RNA hybrids by centrifuge, the
subtracted cDNA libraries were constructed by
transformation of the remaining DNA into competent cells E.
coli Top10F' with electroporation method."
BASE COUNT 75 a 173 c 159 g 82 t
ORIGIN
Query Match 1.9%; Score 32; DB 14; Length 489;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGGATGCGCGCTGCCAGCTGGGACTGCC 685
|||||
Db 139 GAGGATGCGCGCTGCCAGCTGGGACTGCC 170

RESULT 15
BG538695 618 bp mRNA linear EST 03-APR-2001
LOCUS 602566932F1 NIH_MGC_77 Homo sapiens CDNA clone IMAGE:4691574 5',
DEFINITION mRNA sequence.
ACCESSION BG538695
VERSION BG538695.1 GI:13530928
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 (bases 1 to 618)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-rc@mail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM1510 row: p column: 07
 High quality sequence stop: 499.

FEATURES
 source
 1. 618
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4691574"
 /clone_lib="NIH_MGC_77"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
 SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGGCCGAGCGGCCGACATG-dT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.9
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

BASE COUNT 95 a 194 c 213 g 116 t
 ORIGIN

Query Match 1.9%; Score 32; DB 12; Length 618;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGTGCTGATGGCTGTGCTGCTG 406
 ||||||||||||||||||||||||||||||||
 Db 291 GTACCCCTGTGCTGATGGCTGTGCTGCTG 322

Search completed: July.28, 2003, 18:02:02
 Job time : 3040.91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:19:16 ; Search time 486.326 Seconds

(without alignments)
10472.371 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708

Perfect score: 175
Sequence: 1 AGTCACGAGACTTGAGCTTT.....GCCTAGAAATAACACCCCAA 175

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 segs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	58.3	137964	2 AC126895	AC126895 Rattus no
2	102	58.3	226303	2 AC095418	AC095418 Rattus no
3	69	39.4	1741	10 AF259981	AF259981 Rattus no
4	35	20.0	1734	6 AR210324	AR210324 Sequence
5	35	20.0	1734	6 AR210325	AR210325 Sequence
6	35	20.0	1734	10 AF100778	AF100778 Mus muscu
7	35	20.0	61072	10 AL731698	AL731698 Mouse DNA
8	35	20.0	216757	2 AL669906	AL669906 Mus muscu
9	33	18.9	1739	10 AF126063	AF126063 Mus muscu
10	23	13.1	145540	2 AC015962	AC015962 Homo sapi
11	23	13.1	176107	2 AP002393	AP002393 Homo sapi
12	23	13.1	176612	2 AC090246	AC090246 Homo sapi
13	23	13.1	185568	9 AC105227	AC105227 Homo sapi
14	21	12.0	167691	2 AC103349	AC103349 Mus muscu
15	21	12.0	173702	2 AC114696	AC114696 Rattus no
16	21	12.0	231791	2 AL732620	AL732620 Mus muscu
17	20	11.4	25546	2 AC017886	AC017886 Drosophill
18	20	11.4	49759	3 AE002717	AE002717 Drosophill
19	20	11.4	74534	3 AC004735	AC004735 Drosophill
20	20	11.4	102714	9 AP000692	AP000692 Homo sapi
21	20	11.4	106735	3 AC006215	AC006215 Drosophill
22	20	11.4	127724	2 AC129134	AC129134 Rattus no
23	20	11.4	138653	2 AL445431	AL445431 Human DNA
24	20	11.4	148879	2 AC115134	AC115134 Rattus no
25	20	11.4	159119	2 AC120233	AC120233 Rattus no
26	20	11.4	160929	9 AC008677	AC008677 Homo sapi
27	20	11.4	175835	2 AC024074	AC024074 Homo sapi
28	20	11.4	184681	2 AC084409	AC084409 Mus muscu
29	20	11.4	189672	9 CNS01DVW	AL136059 Human chr
30	20	11.4	191895	3 AC093197	AC093197 Drosophill
31	20	11.4	206803	2 AC115932	AC115932 Mus muscu
32	20	11.4	309928	3 AE003666	AE003666 Drosophill
33	20	11.4	340000	9 AP001725	AP001725 Homo sapi
34	20	11.4	60372	2 AC097209	AC097209 Rattus no
35	19	10.9	140726	2 AC122991	AC122991 Rattus no
36	19	10.9	146463	2 AC068306	AC068306 Homo sapi
37	19	10.9	146937	2 AC109876	AC109876 Rattus no
38	19	10.9	160199	9 AF391285	AF391285 Homo sapi
39	19	10.9	163713	2 AC069432	AC069432 Homo sapi
40	19	10.9	173797	2 AC027289	AC027289 Homo sapi
41	19	10.9	178980	10 AL669897	AL669897 Mouse DNA
43	19	10.9	181721	9 AC010879	AC010879 Homo sapi
44	19	10.9	182522	9 AC015689	AC015689 Homo sapi
45	19	10.9	189377	2 AC113164	AC113164 Rattus no

ALIGNMENTS

RESULT 1
AC126895 137964 bp DNA HTG 24-JUL-2002
LOCUS Rattus norvegicus clone CH230-301E4, *** SEQUENCING IN PROGRESS
DEFINITION *** 49 unordered pieces.
ACCESSION AC126895
VERSION AC126895.1 GI:21724040
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 137964)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Pred. No. is the number of results predicted by chance to have a

Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.E., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Eamhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsso, E., Kelly, S., Khan, U., King, L., Kovah, J., Kovar, C.,
Kratoch, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwodu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,
Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 137964)
Worley, K.C.

Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 137964)
Worley, K.C.

Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G2HG
Center clone name: CH230-301E4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye 1.0 of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89474 bases at least Q40
Consensus quality: 93422 bases at least Q30
Consensus quality: 96506 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1056: contig of 1056 bp in length
1057 1156: gap of unknown length
1157 2335: contig of 1179 bp in length
2336 2435: gap of unknown length
2436 3442: contig of 1007 bp in length
3443 3542: gap of unknown length
3543 5081: contig of 1539 bp in length.
5082 5181: gap of unknown length
5182 6424: contig of 1243 bp in length
6425 6524: gap of unknown length
6525 7814: contig of 1290 bp in length
7815 7914: gap of unknown length
7915 9153: contig of 1239 bp in length
9154 9253: gap of unknown length
9254 10445: contig of 1192 bp in length
10446 10545: gap of unknown length
10546 11771: contig of 1226 bp in length
11772 11871: gap of unknown length
11872 13240: contig of 1369 bp in length
13241 13340: gap of unknown length
13341 14359: contig of 1019 bp in length
14360 14459: gap of unknown length
14460 15651: contig of 1192 bp in length
15652 15751: gap of unknown length
15752 17494: contig of 1743 bp in length
17495 17594: gap of unknown length
17595 18679: contig of 1085 bp in length
18680 18779: gap of unknown length
18780 20681: contig of 1902 bp in length
20682 20781: gap of unknown length
20782 22118: contig of 1337 bp in length
22119 22218: gap of unknown length
22219 23578: contig of 1360 bp in length
23579 23678: gap of unknown length
23679 25423: contig of 1745 bp in length
25424 25523: gap of unknown length
25524 27808: contig of 2285 bp in length
27809 27908: gap of unknown length
27909 30272: contig of 2364 bp in length
30273 30372: gap of unknown length
30373 33091: contig of 2719 bp in length
33092 33191: gap of unknown length
33192 35777: contig of 2586 bp in length
35778 35877: gap of unknown length
35879 36908: contig of 1031 bp in length
36909 37008: gap of unknown length
37009 38930: contig of 1922 bp in length
38931 39030: gap of unknown length
39031 41210: contig of 2180 bp in length
41211 41310: gap of unknown length
41311 42937: contig of 1627 bp in length
42938 43037: gap of unknown length
43039 46120: contig of 3083 bp in length
46121 46220: gap of unknown length
46221 48217: contig of 1997 bp in length
48218 48317: gap of unknown length
48319 50433: contig of 2116 bp in length
50434 50533: gap of unknown length
50534 52880: contig of 2347 bp in length
52881 52980: gap of unknown length
52981 55194: contig of 2214 bp in length
55195 55294: gap of unknown length
55295 57482: contig of 2188 bp in length
57483 57582: gap of unknown length
57583 60781: contig of 3199 bp in length
60782 60881: gap of unknown length
60882 62599: contig of 1718 bp in length

* 62600 62699: gap of unknown length
* 62700 65203: contig of 2504 bp in length
* 65204 65303: gap of unknown length
* 65304 66914: contig of 1611 bp in length
* 66915 67014: gap of unknown length
* 67015 70839: contig of 3825 bp in length
* 70840 70939: gap of unknown length
* 70940 75139: contig of 4200 bp in length
* 75140 75239: gap of unknown length
* 75240 78296: contig of 3057 bp in length
* 78297 78396: gap of unknown length
* 78397 83138: contig of 4742 bp in length
* 83139 83238: gap of unknown length
* 83239 88204: contig of 4966 bp in length
* 88205 88304: gap of unknown length
* 88305 92238: contig of 3934 bp in length
* 92239 92338: gap of unknown length
* 92339 97339: contig of 5001 bp in length
* 97340 97439: gap of unknown length
* 97440 103534: contig of 6095 bp in length
* 103535 103634: gap of unknown length
* 103635 107080: contig of 3446 bp in length
* 107081 107180: gap of unknown length
* 107181 115201: contig of 8021 bp in length
* 115202 115301: gap of unknown length
* 115302 121237: contig of 5936 bp in length
* 121238 121337: gap of unknown length
* 121338 127910: contig of 6573 bp in length
* 127911 128010: gap of unknown length
* 128011 137964: contig of 9954 bp in length.
Location/Qualifiers
1. 137964
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-301E4"

BASE COUNT 31831 a 34027 c 35229 g 31979 t 4898 others
ORIGIN

Query Match 58.3%; Score 102; DB 2; Length 137964;
Best Local Similarity 99.3%; Pred. No. 1e-49;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTCAGGAAGCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 44819 AGTCAGGAAGCTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 44878
QY 61 AGGAAGGCTCCACACCTCTGGCAGGCGGACCTTCTCTTCAGCATGAGAAAGACAAGG 120
|||||
Db 44879 AGGAAGGCTCCACACCTCTGGTAGGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAGG 44938
QY 121 GACAGCAGAGTACTCTCTCTGAGAGACTAGTC 153
|||||
Db 44939 GACAGCAGAGTACTCTCTCTGAGAGACTAGTC 44971

RESULT 2
LOCUS AC095418 226303 bp DNA linear HTG 11-JUL-2002
DEFINITION Rattus norvegicus clone CH230-7C10, *** SEQUENCING IN PROGRESS ***,
51 unordered pieces.
AC095418
AC095418.3 GI:21717893
VERSION HTG; HTGS_PHASE1.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 226303)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Edwards-C.C., Elhaj,C., Escotto,M.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Foster,P., Frantz,P.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Garner,T., Garza,N., Gill,R.,
Gabisl,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshitari,N., Stinson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 226303)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 226303)
Worley,K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 10, 2002 this sequence version replaced gi:17941885.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCDF
Center clone name: CH230-7C10
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 185950 bases at least Q40
Consensus quality: 190362 bases at least Q30
Consensus quality: 193076 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 51 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1051:	contig of 1051 bp in length
1052	1151:	gap of unknown length
1152	2286:	contig of 1135 bp in length
2287	2386:	gap of unknown length
2387	3433:	contig of 1047 bp in length
3434	3533:	gap of unknown length
3534	4614:	contig of 1081 bp in length
4615	4714:	gap of unknown length
4715	6581:	contig of 1867 bp in length
6582	6681:	gap of unknown length
6682	7773:	contig of 1092 bp in length
7774	7873:	gap of unknown length
7874	9530:	contig of 1657 bp in length
9531	9630:	gap of unknown length
9631	10661:	contig of 1031 bp in length
10662	10761:	gap of unknown length
10762	11958:	contig of 1197 bp in length
11959	12058:	gap of unknown length
12059	13140:	contig of 1082 bp in length
13141	13240:	gap of unknown length
13241	14777:	contig of 1537 bp in length
14778	14877:	gap of unknown length
14878	16063:	contig of 1186 bp in length
16064	16163:	gap of unknown length
16164	17585:	contig of 1422 bp in length
17586	17685:	gap of unknown length
17686	18916:	contig of 1231 bp in length
18917	19016:	gap of unknown length
19017	20305:	contig of 1289 bp in length
20306	20405:	gap of unknown length
20406	21537:	contig of 1132 bp in length
21538	21637:	gap of unknown length
21638	23703:	contig of 2066 bp in length
23704	23803:	gap of unknown length
23804	26306:	contig of 2503 bp in length
26307	26406:	gap of unknown length
26407	28431:	contig of 2025 bp in length
28432	28531:	gap of unknown length
28532	30615:	contig of 2084 bp in length
30616	30715:	gap of unknown length
30716	33215:	contig of 2500 bp in length
33216	33315:	gap of unknown length
33316	35357:	contig of 2042 bp in length
35358	35457:	gap of unknown length
35458	38012:	contig of 2555 bp in length
38013	38112:	gap of unknown length
38113	40338:	contig of 2226 bp in length
40339	40438:	gap of unknown length
40439	44051:	contig of 3613 bp in length
44052	44151:	gap of unknown length
44152	46294:	contig of 2143 bp in length
46295	46394:	gap of unknown length
46395	48845:	contig of 2451 bp in length
48846	48945:	gap of unknown length
48946	51753:	contig of 2808 bp in length
51754	51853:	gap of unknown length
51854	55915:	contig of 4062 bp in length
55916	56015:	gap of unknown length
56016	60468:	contig of 4453 bp in length
60469	60568:	gap of unknown length
60569	63640:	contig of 3072 bp in length
63641	63740:	gap of unknown length
63741	67888:	contig of 4148 bp in length
67889	67988:	gap of unknown length
67989	73014:	contig of 5026 bp in length
73015	73114:	gap of unknown length
73115	77313:	contig of 4199 bp in length
77314	77413:	gap of unknown length
77414	83786:	contig of 6373 bp in length

83787	83886:	gap of unknown length
83887	89590:	contig of 5704 bp in length
89591	89690:	gap of unknown length
89691	93846:	contig of 4156 bp in length
93847	93946:	gap of unknown length
93947	98126:	contig of 4180 bp in length
98127	98226:	gap of unknown length
98227	104258:	contig of 6032 bp in length
104259	104358:	gap of unknown length
104359	110994:	contig of 6636 bp in length
110995	111094:	gap of unknown length
111095	119838:	contig of 8744 bp in length
119839	119938:	gap of unknown length
119939	126339:	contig of 6401 bp in length
126340	126439:	gap of unknown length
126440	134888:	contig of 8449 bp in length
134889	134988:	gap of unknown length
134989	142709:	contig of 7721 bp in length
142710	142809:	gap of unknown length
142810	149818:	contig of 7009 bp in length
149819	149918:	gap of unknown length
149919	158719:	contig of 8801 bp in length
158720	158819:	gap of unknown length
158820	169868:	contig of 11049 bp in length
169869	169968:	gap of unknown length
169969	180390:	contig of 10422 bp in length
180391	180490:	gap of unknown length
180491	191272:	contig of 10782 bp in length
191273	191373:	gap of unknown length
191373	209667:	contig of 18295 bp in length
209668	209767:	gap of unknown length
209768	226303:	contig of 16536 bp in length.

location/Qualifiers

1. 226303

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

/clone="CH230-7C10"

BASE COUNT 57875 a 53353 c 52917 g 57041 t 5117 others

Query Match 58.3%; Score 102; DB 2; Length 226303;
Best Local Similarity 99.3%; Pred. No. 1e-49;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	AGTCCAGGAACCTTGAGCTTTGTATTTTCAGAAATGCACATCTCTTAAGCACTCGCAAAAC	60
Db	113012	AGTCCAGGAACCTTGAGCTTTGTATTTTCAGAAATGCACATCTCTTAAGCACTCGCAAAAC	113071
QY	61	AGGAAGGCTCCACACCTCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAACAAGG	120
Db	113072	AGGAAGGCTCCACACCTCTGGTAGGCCAGGCGCTTCTCTTCAGCATGAGAACAAGG	113131
QY	121	GACAGCAGAGTACTCTCTCTGAGGAGCTAGTC	153
Db	113132	GACAGCAGAGTACTCTCTCTGAGGAGCTAGTC	113164

RESULT 3

AF259981 1741 bp mRNA linear ROD 09-MAY-2000

LOCUS Rattus norvegicus CCN family protein COP-1 (COP-1) mRNA, complete cds.

ACCESSION AF259981

VERSION AF259981.1 GI:7739780

KEYWORDS Rattus norvegicus.

SOURCE Rattus norvegicus.

ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1741)

AUTHORS Zhang, R., Averboukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P. J., Coffey, R. J., Pardee, A. B. and Liang, P.

TITLE Identification of rCop-1, a new member of the CCN protein family,

JOURNAL as a negative regulator for cell transformation
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)

MEDLINE 98414629

PUBMED 9742130

REFERENCE 2 (bases 1 to 1741)

AUTHORS Liang, P.

TITLE Direct Submission

JOURNAL Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer Center, 649 MRB II, Nashville, TN 37232, USA

FEATURES

source Location/Qualifiers

1..1741

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

1..1741

/gene="Cop-1"

262..1014

/gene="Cop-1"

/note="secreted protein"

/codon_start=1

/product="CCN family protein COP-1"

/protein_id="AAF69011.1"

/db_xref="GI:7739781"

/translation="MRGSPILRLATSLFLCLLSWCAQLCRPTCPWTPPCPGVPLVLDGCGCKVCARRLTESCEHLHVCEPSQGLVCPGAGPGGHGAVCLDEDDGCEVNGRRYLDGETFKPNCRYLCRCDDGGFTCLPLCSEDTLPMSWDCRPKRIOVPGKCEWVCDQGVTPAIQSRSAQGHQLSALVTPASADAPWPMWSTAMGPCSTTCGLIATRVSNQNRFCOLEIQRLCLPRCLARSHSSWNSAF"

BASE COUNT

386 a 491 c 480 g 384 t

ORIGIN

Query Match

Best Local Similarity 39.4%; Score 69; DB 10; Length 1741;

Matches 69; Conservativity 100.0%; Pred. No. 6e-30;

Mismatches 0; Indels 0; Gaps 0;

QY 11 CTGAGCTTGTATTTCAGGATGCACATCTCTTAAGCACTGCAGAAACAGGAGGCTC 70

Db 1557 CTGAGCTTGTATTTCAGGATGCACATCTCTTAAGCACTGCAGAAACAGGAGGCTC 1616

QY 71 CACACCTCT 79

Db 1617 CACACCTCT 1625

RESULT 4

AR210324

LOCUS AR210324 1734 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 17 from patent US 6387657.

ACCESSION AR210324

VERSION AR210324.1 GI:21512525

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1734)

AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,

Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.

TITLE WISP polypeptides and nucleic acids encoding same

JOURNAL Patent: US 6387657-A 17 14-MAY-2002;

FEATURES Location/Qualifiers

source 1..1734

BASE COUNT 355 a 491 c 495 g 393 t

ORIGIN

Query Match 20.0%; Score 35; DB 6; Length 1734;

Best Local Similarity 100.0%; Pred. No. 1.5e-09;

Matches 35; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGA 111

Db 1611 TCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGA 1645

RESULT 5

AR210325/c

LOCUS AR210325 1734 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 18 from patent US 6387657.

ACCESSION AR210325

VERSION AR210325.1 GI:21512526

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1734)

AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,

Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.

TITLE WISP polypeptides and nucleic acids encoding same

JOURNAL Patent: US 6387657-A 18 14-MAY-2002;

FEATURES Location/Qualifiers

source 1..1734

BASE COUNT 393 a 495 c 491 g 355 t

ORIGIN

Query Match 20.0%; Score 35; DB 6; Length 1734;

Best Local Similarity 100.0%; Pred. No. 1.5e-09;

Matches 35; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGA 111

Db 124 TCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGA 90

RESULT 6

AF100778

LOCUS AF100778 1734 bp mRNA linear ROD 17-DEC-1998

DEFINITION Mus musculus connective tissue growth factor related protein WISP-2

(Wisp2) mRNA, complete cds.

ACCESSION AF100778

VERSION AF100778.1 GI:4028578

KEYWORDS

SOURCE Mus musculus.

ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 1734)

AUTHORS Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A.,

Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C.,

Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D.,

Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.

TITLE WISP genes are members of the connective tissue growth factor

family that are up-regulated in wnt-1-transformed cells and

aberrantly expressed in human colon tumors

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)

FEATURES

source 1..1734

BASE COUNT 99061933

ORIGIN

Query Match 20.0%; Score 35; DB 6; Length 1734;

Best Local Similarity 100.0%; Pred. No. 1.5e-09;

Matches 35; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGA 111

Db 1611 TCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGA 1645

RESULT 5

AR210325/c

LOCUS AR210325 1734 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 18 from patent US 6387657.

ACCESSION AR210325

VERSION AR210325.1 GI:21512526

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1734)

AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,

Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.

TITLE WISP polypeptides and nucleic acids encoding same

JOURNAL Patent: US 6387657-A 18 14-MAY-2002;

FEATURES Location/Qualifiers

source 1..1734

BASE COUNT 99061933

ORIGIN

BASE COUNT	355	a	491	c	495	g	393	t	
ORIGIN									
Query Match	20.0%;	Score	35;	DB	10;	Length	1734;		
Best Local Similarity	100.0%;	Pred. No.	1.5e-09;						
Matches	35;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	77	TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA	111						
Db	1611	TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA	1645						
RESULT 7									
AL731698	61072	bp	DNA	linear	ROD	24-MAY-2002			
LOCUS									
DEFINITION	AL731698	Mouse DNA sequence from clone RP23-161B3 on chromosome 2, complete sequence.							
ACCESSION	AL731698								
VERSION	AL731698.10	GI:21214309							
KEYWORDS	HTG.								
SOURCE	house mouse.								
ORGANISM	Mus musculus								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 61072)								
AUTHORS	Walls, J.								
TITLE	Direct Submission								
JOURNAL	Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk On May 25, 2002 this sequence version replaced gi:21213601. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-161B3 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBACe3.6.								
FEATURES.									
source	1. 61072								
	Location/Qualifiers								
	/organism="Mus musculus"								
	/db_xref="taxon:10090"								
	/chromosome="2"								
	/clone="RP23-161B3"								
	/clone_11b="RPCI-23"								
BASE COUNT	15405	a	15368	c	15033	g	15266	t	
ORIGIN									
Query Match	20.0%;	Score	35;	DB	10;	Length	61072;		
Best Local Similarity	100.0%;	Pred. No.	1.6e-09;						
Matches	35;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;

```

OY      77 TCTGCAGGCCAGGCCCTTCTCTCAGCATGAGA 1111
        |||
Db      17823 TCTGCAGGCCAGGCCCTTCTCTCAGCATGAGA 17857

RESULT 8
LOCUS   AL669906
DEFINITION Mus musculus chromosome 2 clone RP23-217C2, *** SEQUENCING IN
ACCESSION AL669906
VERSION   AL669906.5 GI:21955520
KEYWORDS  HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE    Mus musculus
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS  Direct Submission
TITLE     Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL   Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
          humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
          On Jul 25, 2002 this sequence version replaced gi:18181793.

COMMENT  ----- Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: humquerry@sanger.ac.uk
          ----- Project Information
          Center project name: bM217C2
          ----- Summary Statistics
          Assembly program: XGAP4; version 4.5
          Chemistry: Dye-terminator Big Dye; 100% of reads
          Consensus quality: 211865 bases at least Q40
          Consensus quality: 213403 bases at least Q30
          Consensus quality: 214139 bases at least Q20
          Insert size: 214757; sum-of-contigs
          Insert size: 234243; 1.8% error; agarose-fp
          Quality coverage: 6.29x in Q20 bases; sum-of-contigs Quality
          coverage: 5.90x in Q20 bases; agarose-fp
          -----
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 21 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *
          * 1 4000: contig of 4000 bp in length
          * 4001 4100: gap of 100 bp
          * 4101 15267: contig of 11167 bp in length
          * 15268 15367: gap of 100 bp
          * 15368 34970: contig of 19603 bp in length
          * 34971 35070: gap of 100 bp
          * 35071 39671: contig of 4601 bp in length
          * 39672 39771: gap of 100 bp
          * 39772 46360: contig of 6589 bp in length
          * 46361 46460: gap of 100 bp
          * 46461 49809: contig of 3349 bp in length
          * 49810 49909: gap of 100 bp
          * 49910 54509: contig of 4600 bp in length
          * 54510 54609: gap of 100 bp
          * 54610 65989: contig of 11380 bp in length
          * 65990 66089: gap of 100 bp
          * 66090 84635: contig of 18546 bp in length
          * 84636 84735: gap of 100 bp
          * 84736 87399: contig of 2664 bp in length
          * 87400 87499: gap of 100 bp
          * 87500 98601: contig of 11102 bp in length
          * 98602 98701: gap of 100 bp

```

* 98702 105522: contig of 6821 bp in length
* 105523 105622: gap of 100 bp
* 105623 119773: contig of 14151 bp in length
* 119774 119874: gap of 100 bp
* 119874 134552: contig of 14679 bp in length
* 134553 134652: gap of 100 bp
* 134653 141202: contig of 6550 bp in length
* 141203 141302: gap of 100 bp
* 141303 145808: contig of 4506 bp in length
* 145809 145908: gap of 100 bp
* 145909 151572: contig of 5664 bp in length
* 151573 151672: gap of 100 bp
* 151673 169686: contig of 18014 bp in length
* 169687 169786: gap of 100 bp
* 169787 200095: contig of 30309 bp in length
* 200096 200195: gap of 100 bp
* 200196 213877: contig of 13682 bp in length
* 213878 213977: gap of 100 bp
* 213978 216757: contig of 2780 bp in length.

FEATURES
Source Location/Qualifiers
1. 216757

/organism="Mus musculus"
/db_xref="taxon:10090"

/chromosome="2"
/clone="RP23-217C2"

/clone_lib="RPCI-23"
1. 4000

misc_feature /note="assembly_fragment:03429
clone_end:SP6
vector_side:left"

misc_feature /note="assembly_fragment:00814
fragment_chain:1"

misc_feature /note="assembly_fragment:00812
fragment_chain:1"

misc_feature /note="assembly_fragment:03544
fragment_chain:1"

misc_feature /note="assembly_fragment:01438
fragment_chain:1"

misc_feature /note="assembly_fragment:03410
fragment_chain:2"

misc_feature /note="assembly_fragment:00811
fragment_chain:2"

misc_feature /note="assembly_fragment:02217
fragment_chain:2"

misc_feature /note="assembly_fragment:03529
fragment_chain:2"

misc_feature /note="assembly_fragment:01019
fragment_chain:3"

misc_feature /note="assembly_fragment:01170
fragment_chain:3"

misc_feature /note="assembly_fragment:02829
fragment_chain:3"

misc_feature /note="assembly_fragment:01476
fragment_chain:4"

misc_feature /note="assembly_fragment:00505
fragment_chain:4"

misc_feature /note="assembly_fragment:01951
fragment_chain:4"

misc_feature /note="assembly_fragment:01951
fragment_chain:4"

/note="assembly_fragment:02637
fragment_chain:5"
145909. 151572
/note="assembly_fragment:02352
fragment_chain:5"
151673. 169686
/note="assembly_fragment:02207
fragment_chain:5"
169787. 200095
/note="assembly_fragment:02849
fragment_chain:6"
200196. 213877
/note="assembly_fragment:01052
fragment_chain:6"
213978. 216757
/note="assembly_fragment:01318
clone_end:T7
vector_side:right"
BASE COUNT 56425 a 51573 c 51904 g 54852 t 2003 others
ORIGIN

Query Match 20.0%; Score 35; DB 2; Length 216757;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 TCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGA 111
|||||
Db 88929 TCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGA 88963

RESULT 9
AF126063
LOCUS

DEFINITION Mus musculus connective tissue growth factor-like protein precursor
(Ctgfl) mRNA, complete cds.

ACCESSION AF126063
VERSION AF126063.1 GI:4337059

KEYWORDS
SOURCE

ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 1739)
AUTHORS Kumar,S., Hand,A.T., Connor,J.R., Dodds,R.A., Ryan,P.J.,
Trill,J.J., Fisher,S.M., Nuttall,M.E., Lipshutz,D.B., Zou,C.,
Hwang,S.M., Volta,B.J., James,I.E., Rieman,D.J., Gowen,M. and
lee,J.C.

TITLE

Identification and cloning of a connective tissue growth
factor-like cDNA from human osteoblasts encoding a novel regulator
of osteoblast functions

JOURNAL J. Biol. Chem. 274 (24), 17123-17131 (1999)
MEDLINE 99287915
PUBMED 10358067

REFERENCE

AUTHORS Kumar,S. and Zou,C.

TITLE Direct Submission

JOURNAL Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109,
Smithkline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406,
USA

FEATURES
Source Location/Qualifiers
1. 1739

/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="lung"

gene 1. 1739
/gene="Ctgfl"

CDS 242. 997
/gene="Ctgfl"

/note="similar to the Mus musculus WISP-2 protein encoded
by the sequence presented in GenBank Accession Number
AF100778; putative growth factor; CTGF-L; contains IGF
binding (IGFBD), Von Willebrand Factor type C (WVC) repeat
and thrombospondin type I (TSP1) domains; member of the

CCN (CTGF/Cyrl61/Nov) family; lacks the fourth
carboxy-terminal (CT) domain present in other members of
the CCN family"
/codon_start=1
/product="connective tissue growth factor-like protein
precursor"
/protein_id="AADI8058.1"
/db_xref="GI:4337060"
/translation="MRGNPLIHLLAISFLCILSNVYSQLCPAPCACPWTPPQCPGPV
LVLDGCGCCRCVRCARLSCDHLHYCDPSQGLVQPGAGPSGRGAVCLFEEDGSCVE
NGRRYLDGETFKPNCRVLCRCDDGGTCLPLCSEDEVRLPSWDCPPRRRIQVGRCCPE
WVCDQAVMQPAIQSSAOGHQLSALVTPASADGPCPMWSTAWGPCSTTCGLGIATRV
NONRFGQLEIQRLCLSRPCLASRSHGWSNSAE"

BASE COUNT 375 a 480 c 489 g 395 t

ORIGIN

Query Match 18.9%; Score 33; DB 10; Length 1739;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 TGGCAGGCGAGGCGCTTCTCTCAGCATGAGA 111
|||||
Db 1598 TGGCAGGCGAGGCGCTTCTCTCAGCATGAGA 1630

RESULT 10
AC015962 145540 bp DNA linear HTG 09-MAY-2001
LOCUS Homo sapiens chromosome 18 clone RP11-111D6 map 18, WORKING DRAFT
DEFINITION
AC015962
AC015962
AC015962
AC015962.4 GI:7382674
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 145540)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone RP11-111D6
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 145540)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barina,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearrellano,K., Dewar,K., Doming,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hages,B., Heaford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Margulis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tlirell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 1, 2000 this sequence version replaced gi:6957735.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L970
Center clone name: 111_D_6
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-primer-amersham; 4% of reads
Chemistry: Dye-terminator Big Dye; 96% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 125662 bases at least Q40
Consensus quality: 134718 bases at least Q30
Consensus quality: 138727 bases at least Q20
Insert size: 80000; agarose-fp
Insert size: 141840; sum-of-contigs
Quality coverage: 5.5 in Q20 bases; agarose-fp
Quality coverage: 3.1 in Q20 bases.
NOTE: This is a 'working draft' sequence. It currently
consists of 38 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1060: contig of 1060 bp in length
1061 1160: gap of 100 bp
1161 2171: contig of 1011 bp in length
2172 2271: gap of 100 bp
2272 3611: contig of 1340 bp in length
3612 3711: gap of 100 bp
3712 5307: contig of 1596 bp in length
5308 5407: gap of 100 bp
5408 7422: contig of 2015 bp in length
7423 7522: gap of 100 bp
7523 9048: contig of 1526 bp in length
9049 9148: gap of 100 bp
9149 10173: contig of 1025 bp in length
10174 10273: gap of 100 bp
10274 11773: contig of 1500 bp in length
11774 11873: gap of 100 bp
11874 13476: contig of 1603 bp in length
13477 13576: gap of 100 bp
13577 15396: contig of 1820 bp in length
15397 15496: gap of 100 bp
15497 16981: contig of 1485 bp in length
16982 17081: gap of 100 bp
17082 19160: contig of 2079 bp in length
19161 19260: gap of 100 bp
19261 21188: contig of 1928 bp in length
21189 21288: gap of 100 bp
21289 23565: contig of 2277 bp in length
23566 23665: gap of 100 bp
23666 25713: contig of 2048 bp in length
25714 25813: gap of 100 bp
25814 26916: contig of 1103 bp in length
26917 27016: gap of 100 bp
27017 29802: contig of 2786 bp in length
29803 29902: gap of 100 bp
29903 32825: contig of 2923 bp in length
32826 32925: gap of 100 bp
32926 35752: contig of 2827 bp in length
35753 35852: gap of 100 bp
35853 37965: contig of 2113 bp in length
37966 38065: gap of 100 bp
38066 41361: contig of 3296 bp in length
41362 41461: gap of 100 bp
41462 44463: contig of 3002 bp in length
44464 44563: gap of 100 bp
44564 47687: contig of 3124 bp in length
47688 47787: gap of 100 bp
47788 51777: contig of 3990 bp in length
51778 51877: gap of 100 bp
51878 56131: contig of 4254 bp in length
56132 56231: gap of 100 bp
56232 60894: contig of 4663 bp in length
60895 60994: gap of 100 bp
60995 65362: contig of 4368 bp in length
65363 65462: gap of 100 bp
65463 69104: contig of 3642 bp in length

* 69105 69204: gap of 100 bp
* 69205 73085: contig of 3881 bp in length
* 73086 73185: gap of 100 bp
* 73186 77580: contig of 4395 bp in length
* 77581 77680: gap of 100 bp
* 77681 85326: contig of 7646 bp in length
* 85327 85426: gap of 100 bp
* 85427 92750: contig of 7324 bp in length
* 92751 92850: gap of 100 bp
* 92851 100690: contig of 7840 bp in length
* 100691 100790: gap of 100 bp
* 100791 108177: contig of 7387 bp in length
* 108178 108277: gap of 100 bp
* 108278 115728: contig of 7451 bp in length
* 115729 115828: gap of 100 bp
* 115829 125137: contig of 9309 bp in length
* 125138 125237: gap of 100 bp
* 125238 134119: contig of 8882 bp in length
* 134120 134219: gap of 100 bp
* 134220 145540: contig of 11321 bp in length.

FEATURES

SOURCE

1. 145540
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-111D6"
/clone_lib="RPC1-11 Human Male BAC"
1. 1060
/note="assembly_fragment"
1161. 2171
/note="assembly_fragment"
2272. 3611
/note="assembly_fragment"
3712. 5307
/note="assembly_fragment"
5408. 7422
/note="assembly_fragment"
7523. 9048
/note="assembly_fragment"
9149. 10173
/note="assembly_fragment"
10274. 11773
/note="assembly_fragment"
11874. 13476
/note="assembly_fragment"
13577. 15396
/note="assembly_fragment"
15497. 16981
/note="assembly_fragment"
17082. 19160
/note="assembly_fragment"
19261. 21188
/note="assembly_fragment"
21289. 23565
/note="assembly_fragment"
23666. 25713
/note="assembly_fragment"
25814. 26916
/note="assembly_fragment"
clone_end:SP6
vector_side:left
27017. 29802
/note="assembly_fragment"
29903. 32825
/note="assembly_fragment"
32926. 35752
/note="assembly_fragment"
35853. 37965
/note="assembly_fragment"
38066. 41361
/note="assembly_fragment"
41462. 44463

/note="assembly_fragment"
44564. 47687
/note="assembly_fragment"
47788. 51777
/note="assembly_fragment"
51878. 56131
/note="assembly_fragment"
56232. 60894

Query Match 13.1%; Score 23; DB 2; Length 145540;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 CTTGAGCTTTGTATTTTCAGGAA 33
|||||
Db 87756 CTTGAGCTTTGTATTTTCAGGAA 87778

RESULT 11

AP002393/c

LOCUS

AP002393 176107 bp DNA linear HTG 31-MAY-2000
Homo sapiens chromosome 18 clone RP11-699C17 map 18q12, WORKING

DEFINITION
DRAFT SEQUENCE, 19 unordered pieces.

ACCESSION
AP002393

VERSION
AP002393.1 GI:8131657

KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE
Homo sapiens DNA, clone:RP11-699C17.

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 176107)

AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 176,107 genomic DNA of 18q12

Published Only in Database (2000)

2 (bases 1 to 176107)

AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Direct Submission

Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,

Japan (E-mail:hattori@gsc.riken.go.jp,

URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,

Fax:81-42-778-9924)

----- Genome Center

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

----- Project Information

Center project name: RP11-699C17

Center clone name: RP11-699C17

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 163826 bases at least Q40

Consensus quality: 169603 bases at least Q30

Consensus quality: 172435 bases at least Q20

Insert size: 174307; sum-of-contigs

Quality coverage: 5.63x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of

19 contigs. The true order of the pieces is not known and their

order in this sequence record is arbitrary. Gaps between the

contigs are represented as runs N, but the exact sizes of the gaps

are unknown. This record will be updated with the finished sequence

as soon as it is available and the accession number will be

preserved

1 31231 contig of 31231 bp in length
31332 61792 contig of 30461 bp in length
61893 88918 contig of 27026 bp in length


```

89019 98085 contig of 9067 bp in length
98186 109413 contig of 11228 bp in length
109514 119559 contig of 10046 bp in length
119660 128623 contig of 8964 bp in length
128724 136240 contig of 7517 bp in length
136341 144191 contig of 7851 bp in length
144292 149956 contig of 5665 bp in length
150057 156749 contig of 6693 bp in length
156850 162315 contig of 5466 bp in length
162416 164672 contig of 2257 bp in length
164773 167491 contig of 2719 bp in length
167592 170792 contig of 3201 bp in length
170893 172798 contig of 1906 bp in length
172899 173274 contig of 376 bp in length
173375 174909 contig of 1535 bp in length
174910 175009 gap of 100 bp
175010 176107 contig of 1098 bp in length.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 31231: contig of 31231 bp in length
31232 31331: gap of 100 bp
31332 61792: contig of 30461 bp in length
61793 61892: gap of 100 bp
61893 88918: contig of 27026 bp in length
88919 89018: gap of 100 bp
89019 98085: contig of 9067 bp in length
98086 98185: gap of 100 bp
98186 109413: contig of 11228 bp in length
109414 109513: gap of 100 bp
109514 119559: contig of 10046 bp in length
119560 119659: gap of 100 bp
119660 128623: contig of 8964 bp in length
128624 128723: gap of 100 bp
128724 136240: contig of 7517 bp in length
136241 136340: gap of 100 bp
136341 144191: contig of 7851 bp in length
144192 144291: gap of 100 bp
144292 149956: contig of 5665 bp in length
149957 150056: gap of 100 bp
150057 156749: contig of 6693 bp in length
156750 156849: gap of 100 bp
156850 162315: contig of 5466 bp in length
162316 162415: gap of 100 bp
162416 164672: contig of 2257 bp in length
164673 164772: gap of 100 bp
164773 167491: contig of 2719 bp in length
167492 167591: gap of 100 bp
167592 170792: contig of 3201 bp in length
170793 170892: gap of 100 bp
170893 172798: contig of 1906 bp in length
172799 172898: gap of 100 bp
172899 173274: contig of 376 bp in length
173275 173374: gap of 100 bp
173375 174909: contig of 1535 bp in length
174910 175009: gap of 100 bp
175010 176107: contig of 1098 bp in length.

FEATURES
source
1. 176107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q12"
/clone="RP11-699C17"
1. 31231
/misc_feature
/note="assembly_fragment"
31332. 61792
/misc_feature
/note="assembly_fragment"
61893. 88918

```

```

misc_feature /note="assembly_fragment"
89019. 98085
/misc_feature /note="assembly_fragment"
98186. 109413
/misc_feature /note="assembly_fragment"
109514. 119559
/misc_feature /note="assembly_fragment"
119660. 128623
/misc_feature /note="assembly_fragment"
128724. 136240
/misc_feature /note="assembly_fragment"
136341. 144191
/misc_feature /note="assembly_fragment"
144292. 149956
/misc_feature /note="assembly_fragment"
150057. 156749
/misc_feature /note="assembly_fragment"
156850. 162315
/misc_feature /note="assembly_fragment"
162416. 164672
/misc_feature /note="assembly_fragment"
164773. 167491
/misc_feature /note="assembly_fragment"
167592. 170792
/misc_feature /note="assembly_fragment"
170893. 172798
/misc_feature /note="assembly_fragment"
172899. 173274
/misc_feature /note="assembly_fragment clone_end:SP6 vector_side:right"
173375. 174909
/misc_feature /note="assembly_fragment"
175010. 176107
/misc_feature /note="assembly_fragment"
BASE COUNT 46994 a 37813 c 39266 g 50228 t 1806 others
ORIGIN

```

```

Query Match 13.1%; Score 23; DB 2; Length 176107;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 11 CTGAGCTTGTATTTCAGGAA 33
Db 34317 CTGAGCTTGTATTTCAGGAA 34295

```

```

RESULT 12
AC090246/c 176612 bp DNA linear HTG 21-AUG-2002
LOCUS Homo sapiens chromosome 18 clone RP11-699C17 map 18, *** SEQUENCING
DEFINITION IN PROGRESS *** 2 unordered pieces.
ACCESSION AC090246
VERSION AC090246.7 GI:22381545
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 176612)
AUTHORS Birren,B., Nussbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone RP11-699C17
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 176612)
AUTHORS Birren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
Mcpheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

```

Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travls,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 176612)

Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:22123591.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center Project name: LI2328
Center clone name: 699_C_17

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 131563: contig of 131563 bp in length
* 131564 131663: gap of 100 bp
* 131664 176612: contig of 44949 bp in length.

FEATURES
source

1. 176612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-699C17"
/clone_lib="RPC1-11 Human Male BAC"
BASE COUNT 46327 a 37044 c 41351 g 51790 t 100 others
ORIGIN

Query Match 13.1%; Score 23; DB 2; Length 176612;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 CTTGAGCTTTGTATTTTCAGGAA 33
|||||
Db 28978 CTTGAGCTTTGTATTTTCAGGAA 28956

RESULT 13
AC105227/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC105227 185568 bp DNA linear PRI 29-APR-2002
Homo sapiens chromosome 18, clone RP11-1030E3, complete sequence.
AC105227
AC105227.6 GI:20258546
HTG.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lehoczký,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (26-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 185568)

Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Landers,T., Lehoczký,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Direct Submission
Submitted (21-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 4 (bases 1 to 185568)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
Landers,T., Lehoczký,J., Levine,R., Linblad-roh,K., Liu,G.,
Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Mejdridim,J., Menus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tresfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
JOURNAL Submitted (29-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 21, 2002 this sequence version replaced gi:20128372.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23094
Center clone name: 1030_E_3

FEATURES
Source Location/Qualifiers
1.185568
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-1030E3"
/clone_lib="RPC1-11 Human Male PAC"
1042.1066
/rpt_family="AT_rich"
1445.1486
/rpt_family="(TTTA)n"
complement(1541..1793)
/rpt_family="Alusx"
complement(1995..2286)
/rpt_family="LTR4"
complement(2373..2429)
/rpt_family="LTR4"
complement(2616..2927)
/rpt_family="AluY"
complement(2928..3238)
/rpt_family="Alusx"
3557..3648
/rpt_family="L2"
complement(3649..4052)
/rpt_family="MSTB"
4053..4283
/rpt_family="L2"
5396..5724
/rpt_family="MER2"
5964..6103
/rpt_family="LIME3"
6118..7124
/rpt_family="L1PA12"

repeat_region complement(7125..7415)
/rpt_family="AluY"
repeat_region 7416..7577
/rpt_family="L1PA12"
repeat_region 7578..8204
/rpt_family="L1PA14"
repeat_region 8215..8449
/rpt_family="L1ME3"
repeat_region 8671..9139
/rpt_family="L1ME3"
repeat_region 9140..9171
/rpt_family="L1ME3"
repeat_region 9172..10748
/rpt_family="(TTC)n"
repeat_region 10749..12165
/rpt_family="SVA"
repeat_region 12152..12299
/rpt_family="L1ME3"
repeat_region 12303..12481
/rpt_family="L1ME3"
repeat_region 12502..12604
/rpt_family="L1MD1"
repeat_region 13063..13094
/rpt_family="L1ME3"
repeat_region 13063..13094
/rpt_family="AT_rich"
repeat_region complement(13333..14104)
/rpt_family="L1PA7"
repeat_region complement(14164..14291)
/rpt_family="MIR3"
repeat_region complement(14292..14525)
/rpt_family="L2"
repeat_region 14526..15051
/rpt_family="MER1A"
repeat_region complement(15052..15536)
/rpt_family="L2"
repeat_region complement(15535..15608)
/rpt_family="MIR"
repeat_region 15988..16291
/rpt_family="Alusx"
repeat_region complement(16697..17281)
/rpt_family="L1MC5"
repeat_region 17282..17344
/rpt_family="L1MC5"
repeat_region 17312..17675
/rpt_family="L1MC5"
repeat_region complement(17709..17768)
/rpt_family="L1MC4a"
repeat_region complement(17790..18317)
/rpt_family="MLT1G1"
repeat_region complement(18408..18630)
/rpt_family="L1MC5"
repeat_region complement(18774..18891)
/rpt_family="L1MC5"
repeat_region complement(18899..19462)
/rpt_family="L1MC5"
repeat_region complement(19463..19751)
/rpt_family="L1MC5"
repeat_region complement(19743..22357)
/rpt_family="HERVS71"
repeat_region complement(22351..22457)
/rpt_family="HERVS71"
repeat_region 22458..23434
/rpt_family="L1MC5"

Query Match 13.1%; Score 23; DB 9; Length 185568;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 CTGAGCTTGTATTTTCAGGAA 33
|||||
Db 164413 CTGAGCTTGTATTTTCAGGAA 164391

RESULT 14

AC103349/c
LOCUS 167691 bp DNA linear HTG 21-AUG-2002
DEFINITION Mus musculus clone RP23-303K3, WORKING DRAFT SEQUENCE, 19 unordered
pieces.
AC103349
AC103349.2 GI:22381452
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 167691)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-303K3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 167691)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Larocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Mcpheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (24-NOV-2001) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 167691)
AUTHORS Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
COMMENT Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:17063224.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information

Center project name: L17657
Center clone name: 303_K_3
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 161237 bases at least Q40
Consensus quality: 164563 bases at least Q30
Consensus quality: 165460 bases at least Q20
Insert size: 163000; agarose-fp
Insert size: 165891; sum-of-contigs
Quality coverage: 8.8 in Q20 bases; agarose-fp
Quality coverage: 8.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4555: contig of 4555 bp in length
* 4556 4655: gap of 100 bp
* 4656 5264: contig of 609 bp in length
* 5265 5364: gap of 100 bp
* 5365 5944: contig of 580 bp in length
* 5945 6044: gap of 100 bp
* 6045 7254: contig of 1210 bp in length
* 7255 7354: gap of 100 bp
* 7355 8690: contig of 1336 bp in length
* 8691 8790: gap of 100 bp
* 8791 9904: contig of 1114 bp in length
* 9905 10004: gap of 100 bp
* 10005 11421: contig of 1417 bp in length
* 11422 11521: gap of 100 bp
* 11522 12995: contig of 1474 bp in length
* 12996 13095: gap of 100 bp
* 13096 14365: contig of 1270 bp in length
* 14366 14465: gap of 100 bp
* 14466 16284: contig of 1819 bp in length
* 16285 16384: gap of 100 bp
* 16385 18855: contig of 2471 bp in length
* 18856 18955: gap of 100 bp
* 18956 22999: contig of 4044 bp in length
* 23000 23099: gap of 100 bp
* 23100 26210: contig of 3111 bp in length
* 26211 26310: gap of 100 bp
* 26311 40871: contig of 14561 bp in length
* 40872 40971: gap of 100 bp
* 40972 56546: contig of 15575 bp in length
* 56547 56646: gap of 100 bp
* 56647 80881: contig of 24235 bp in length
* 80882 80981: gap of 100 bp
* 80982 103267: contig of 22286 bp in length
* 103268 103367: gap of 100 bp
* 103368 126955: contig of 23588 bp in length
* 126956 127055: gap of 100 bp
* 127056 167691: contig of 40636 bp in length.
FEATURES
source
1..167691
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-303K3"
/clone_lib="RPCI-23 Female Mouse BAC"
1..4555
/note="assembly_fragment
clone_end:SP6
vector_side:left"
4656..5264
/note="assembly_fragment"
5365..5944
/note="assembly_fragment"


```

misc_feature 6045. 7254
                /note="assembly_fragment"
misc_feature 7355. 8690
                /note="assembly_fragment"
misc_feature 8791. 9904
                /note="assembly_fragment"
misc_feature 10005. 11421
                /note="assembly_fragment"
misc_feature 11522. 12995
                /note="assembly_fragment"
misc_feature 13096. 14365
                /note="assembly_fragment"
misc_feature 14466. 16284
                /note="assembly_fragment"
misc_feature 16385. 18855
                /note="assembly_fragment"
misc_feature 18956. 22999
                /note="assembly_fragment"
misc_feature 23100. 26210
                /note="assembly_fragment"
misc_feature 26311. 40871
                /note="assembly_fragment"
misc_feature 40972. 56546
                /note="assembly_fragment"
misc_feature 56647. 80881
                /note="assembly_fragment"
misc_feature 80982. 103267
                /note="assembly_fragment"
misc_feature 103368. 126955
                /note="assembly_fragment"
misc_feature 127056. 167691
                /note="assembly_fragment"
BASE COUNT 47612 a 35027 c 35914 g 47337 t 1801 others
ORIGIN
Query Match 12.0%; Score 21; DB 2; Length 167691;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 98 TCCTCAGCATGAGAAAGACAA 118
Db 162570 TCCTCAGCATGAGAAAGACAA 162550

RESULT 15
AC114696 173702 bp DNA linear HTG 13-JUL-2002
LOCUS Rattus norvegicus clone CH230-202L22, *** SEQUENCING IN PROGRESS
DEFINITION *** 71 unordered pieces.
AC114696
VERSION AC114696.2 GI:21737289
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 173702)
AUTHORS Muzny,D.M., Adams,C., Adlo-Oduola,B., Alt-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,J., Blincoe,K., Blankenburg,K., Bonnin,D.,
Bouckaert,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

```

```

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsl,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssohn,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenko,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 173702)
Worley,K.C.
Direct Submission
Submitted (11-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:19339033.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOWB
Center clone name: CH230-202L22
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 113417 bases at least Q40
Consensus quality: 117596 bases at least Q30
Consensus quality: 120896 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1324: contig of 1324 bp in length
* 1325 1424: gap of unknown length
* 1425 2953: contig of 1529 bp in length
* 2954 3053: gap of unknown length
* 3054 4276: contig of 1223 bp in length
* 4277 4376: gap of unknown length
* 4377 5562: contig of 1186 bp in length

```

```
* 5563 5662: gap of unknown length
* 5663 6814: contig of 1152 bp in length
* 6815 6914: gap of unknown length
* 6915 8228: contig of 1314 bp in length
* 8229 8328: gap of unknown length
* 8329 9362: contig of 1034 bp in length
* 9363 9462: gap of unknown length
* 9463 11208: contig of 1746 bp in length
* 11209 11308: gap of unknown length
* 11309 12347: contig of 1039 bp in length
* 12348 12447: gap of unknown length
* 12448 13738: contig of 1291 bp in length
* 13739 13838: gap of unknown length
* 13839 14885: contig of 1047 bp in length
* 14886 14985: gap of unknown length
* 14986 16091: contig of 1106 bp in length
* 16092 16191: gap of unknown length
* 16192 17902: contig of 1711 bp in length
* 17903 18002: gap of unknown length
* 18003 19382: contig of 1380 bp in length
* 19383 19482: gap of unknown length
* 19483 20663: contig of 1181 bp in length
* 20664 20763: gap of unknown length
* 20764 21932: contig of 1169 bp in length
* 21933 22032: gap of unknown length
* 22033 23262: contig of 1230 bp in length
* 23263 23362: gap of unknown length
* 23363 25111: contig of 1749 bp in length
* 25112 25211: gap of unknown length
* 25212 26281: contig of 1070 bp in length
* 26282 26381: gap of unknown length
* 26382 27785: contig of 1404 bp in length
* 27786 27885: gap of unknown length
* 27886 29448: contig of 1563 bp in length
* 29449 29548: gap of unknown length
* 29549 31328: contig of 1780 bp in length
* 31329 31428: gap of unknown length
* 31429 33139: contig of 1711 bp in length
* 33140 33239: gap of unknown length
* 33240 34718: contig of 1479 bp in length
* 34719 34818: gap of unknown length
* 34819 36911: contig of 2093 bp in length
* 36912 37011: gap of unknown length
* 37012 38027: contig of 1016 bp in length
* 38028 38127: gap of unknown length
* 38128 39562: contig of 1435 bp in length
* 39563 39662: gap of unknown length
* 39663 40751: contig of 1089 bp in length
* 40752 40851: gap of unknown length
* 40852 42914: contig of 2063 bp in length
* 42915 43014: gap of unknown length
* 43015 44102: contig of 1088 bp in length
* 44103 44202: gap of unknown length
* 44203 45433: contig of 1231 bp in length
* 45434 45533: gap of unknown length
* 45534 46810: contig of 1277 bp in length
* 46811 46910: gap of unknown length
* 46912 48511: contig of 1601 bp in length
* 48512 48611: gap of unknown length
* 48612 50317: contig of 1706 bp in length
* 50318 50417: gap of unknown length
* 50418 52272: contig of 1855 bp in length
* 52273 52372: gap of unknown length
* 52373 53837: contig of 1465 bp in length
* 53838 53937: gap of unknown length
* 53938 55583: contig of 1646 bp in length
* 55584 57128: contig of 1445 bp in length
* 57129 57228: gap of unknown length
* 57229 58882: contig of 1654 bp in length
* 58883 58982: gap of unknown length
* 58983 61320: contig of 2338 bp in length
* 61321 61420: gap of unknown length
```

```
* 61421 63901: contig of 2481 bp in length
* 63902 64001: gap of unknown length
* 64002 66344: contig of 2343 bp in length
* 66345 66444: gap of unknown length
* 66445 67984: contig of 1540 bp in length
* 67985 68084: gap of unknown length
* 68085 70337: contig of 2253 bp in length
* 70338 70437: gap of unknown length
* 70438 72788: contig of 2351 bp in length
* 72789 72888: gap of unknown length
* 72889 74331: contig of 1443 bp in length
* 74332 74432: gap of unknown length
* 74433 77106: contig of 2675 bp in length
* 77107 77206: gap of unknown length
* 77207 79143: contig of 1937 bp in length
* 79144 79243: gap of unknown length
* 79244 82097: contig of 2854 bp in length
* 82098 82197: gap of unknown length
* 82198 85184: contig of 2987 bp in length
* 85185 85284: gap of unknown length
* 85285 87051: contig of 1767 bp in length
* 87052 87151: gap of unknown length
* 87152 89984: contig of 2833 bp in length
* 89985 90084: gap of unknown length
* 90085 91616: contig of 1532 bp in length
* 91617 91716: gap of unknown length
* 91717 93274: contig of 1558 bp in length
```

Query Match 12.0%; Score 21; DB 2; Length 173702;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GCAGGCCAGGGCCTTCTCTT 101
Db 147064 GCAGGCCAGGGCCTTCTCTT 147084

Search completed: July 28, 2003, 21:14:48
Job time : 491.326 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:31:11 ; Search time 45.9617 Seconds
(without alignments)
8574.520 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708
Perfect score: 175
Sequence: 1 AGTCCAGAGACTTGAGCTTT.....GCCTAGATAAACACCCCAA 175

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 segs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	175	100.0	1708	20	AAZ07516
2	121	69.1	439	24	ABL59575
3	54	30.9	65	24	ABN30189
4	35	20.0	1734	20	AAZ76488
5	18	10.3	1668	24	AAZ32413
6	18	10.3	1671	24	ABL91207
7	18	10.3	1671	24	ABL92605
8	18	10.3	2031	24	ABL91770
9	18	10.3	2104	21	AAF21412

C	10	18	10.3	2104	21	AAA35290	Human adenosine re
C	11	18	10.3	2104	22	AAC84222	Human protein kina
C	12	18	10.3	2111	21	AAA75901	DNA encoding a 60
C	13	18	10.3	2586	23	AAS56353	Salmonella typhi D
C	14	18	10.3	5095	22	AAH57492	Human liver cell s
C	15	18	10.3	38644	21	AAF21424	Human low adenosin
C	16	18	10.3	38644	21	AAA35302	Human adenosine re
C	17	18	10.3	1230025	20	AAZ91990	Nucleotide sequenc
C	18	17	9.7	24	20	AAZ76528	Mouse WISP-2 PCR p
C	19	17	9.7	317	21	AAC53400	Arabidopsis thalia
C	20	17	9.7	387	22	AAI80016	Human polynucleoti
C	21	17	9.7	700	22	AAH92347	Human inflammatory
C	22	17	9.7	700	22	AAH92348	Human signal pepti
C	23	17	9.7	742	20	AAZ61274	Human neuroblastom
C	24	17	9.7	791	22	AAI96326	Polypeptide-dihydr
C	25	17	9.7	1026	24	ABL42428	Human MDR1 encodin
C	26	17	9.7	1319	24	ABO72690	Drosophila melanog
C	27	17	9.7	1485	23	ABL05473	Arabidopsis thalia
C	28	17	9.7	1521	21	AAC33310	Human DNA encoding
C	29	17	9.7	2126	22	AAS46172	Human membrane spa
C	30	17	9.7	2261	20	AAZ09837	Human Jurkat cell
C	31	17	9.7	2998	21	AAZ38856	Human Jurkat cell
C	32	17	9.7	2998	21	AAZ38864	Drosophila melanog
C	33	17	9.7	3775	23	ABL05472	Human nervous syst
C	34	17	9.7	6643	22	ABA16216	Mechanical stress
C	35	17	9.7	10427	21	AAZ36325	Human immune/haema
C	36	17	9.7	10672	23	ABL27119	Genomic sequence #
C	37	17	9.7	1761	22	AAK79415	Drosophila melanog
C	38	17	9.7	1761	23	ABK42709	Human transporter
C	39	17	9.7	66743	23	ABL27118	Human tyrosinase
C	40	17	9.7	66804	24	ABK87050	Pyrococcus abyssi
C	41	17	9.7	34980	22	AAF86431	Human silent SNP c
C	42	16	9.1	51	22	AAI76626	Probe #23874 used
C	43	16	9.1	108	22	AAI55188	Human cardiovascular
C	44	16	9.1	219	22	AAS35423	Ovary cancer relat
C	45	16	9.1	229	24	ABL67914	

ALIGNMENTS

RESULT 1	
AAZ07516	
ID	AAZ07516 standard; cDNA; 1708 BP.
XX	
AC	AAZ07516;
XX	
DT	26-NOV-1999 (first entry)
XX	
DE	Rat HICP polypeptide encoding cDNA.
XX	
KW	Heparin-induced CGN-like protein; HICP; cell-associated activity; ss;
KW	cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX	
OS	Rattus sp.
XX	
PN	W09947556-A2.
XX	
PD	23-SEP-1999.
XX	
PF	18-MAR-1999; 99WO-US05999.
XX	
PR	19-MAR-1998; 98US-0044273.
XX	
PA	(TUFT) TUFTS COLLEGE.
XX	
PI	Castellot JJ;
XX	
DR	WPI; 1999-562060/47.
DR	P-PSDB; AAY27434.
XX	
PT	Nucleic acid sequences encoding rat heparin-induced CGN-like protein, used in methods to identify modulators or in diagnostic applications
PT	

XX Claim 2; Fig 1; 108bp; English.
PS
XX This cDNA encodes a rat heparin-induced CN-like protein (HICP) protein.
CC Agents that stimulate or inhibit HICP protein activity or expression,
CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to
CC modulate cell-associated activity. HICP modulators can be used to treat
CC disorders characterized by aberrant HICP protein activity or expression.
CC Probes capable of hybridizing to HICP mRNA or antibodies specific for
CC HICP can be used to detect HICP activity in a biological sample. HICP
CC can be used to treat disorders, such as a cardiovascular or fibrotic
CC disorder, characterized by aberrant cell proliferation.
XX
SQ Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Query Match 100.0%; Score 175; DB 20; Length 1708;
Best Local Similarity 100.0%; Pred. No. 3e-79;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCAGGAAGCTTGTGATTTTTCAGGAATGACATCTCTTAAGCACTCGCAAAAC 60
Db 1534 AGTCAGGAAGCTTGTGATTTTTCAGGAATGACATCTCTTAAGCACTCGCAAAAC 1593

QY 61 AGGAAGGCTCCACACCTCTGCGAGGCGGCTTCTCTTCAGCATGAGAAAGCAAGG 120
Db 1594 AGGAAGGCTCCACACCTCTGCGAGGCGGCTTCTCTTCAGCATGAGAAAGCAAGG 1653

QY 121 GACAGCAGAGTACTCTCTCTGAGAGACTAGTCTAGCCTTGAATAACACCCAA 175
Db 1654 GACAGCAGAGTACTCTCTCTGAGAGACTAGTCTAGCCTTGAATAACACCCAA 1708

RESULT 2
ABL59575
ID ABL59575 standard; DNA; 439 BP.
XX
AC ABL59575;
XX
DT 17-JUL-2002 (first entry)
XX
DE Rat OST23 gene fragment SEQ ID NO:23.
XX
KW Rat; OST; osteoregenerative; parathyroid hormone; tibiae; osteopathic;
KW gene therapy; parathyroid hormone receptor ligand; bone disorder;
KW bone formation disorder; bone resorption disorder; osteoporosis;
KW osteopenia; osteopetrosis; gene; ds.
XX
OS Rattus sp.
XX
PN WO200224943-A2.
PN
PD 28-MAR-2002.
PD
XX
PF 19-SEP-2001; 2001WO-US29548.
PF
XX
PR 19-SEP-2000; 2000US-233579P.
PR
XX
PA (CURA-) CURAGEN CORP.
PA (GLAX) GLAXO GROUP LTD.
PA
XX
PI Horesovsky GJ, Noel LS, Raha D;
PI
XX
DR WPI; 2002-401989/43.
DR
XX
PT Identifying parathyroid hormone receptor ligands and osteoregenerative
PT agents involves detecting the expression of nucleic acids which are
PT regulated by parathyroid hormone
PT
XX
PS Claim 45; Page 22; 90pp; English.
PS
XX The present invention describes a method (M1) for identifying parathyroid
CC hormone receptor ligands (I) and osteoregenerative agents by contacting a
CC test cell population (CP) comprising cells expressing nucleic acid

CC sequences (S) of OST1-47 and 48, with a test agent, measuring nucleic
CC acid sequence expression, comparing it with reference CP and identifying
CC the ligand and agent by the difference in expression levels. Also
CC described is a method (M2) for treating a bone disorder in a subject by
CC administering to the subject an agent that modulates the expression or
CC activity of (S). (I) have osteopathic activities, and can be used to
CC modulate the expression of OST 1-48, and can also be used in gene
CC therapy. (M1) is useful for identifying parathyroid hormone receptor
CC ligands (I) or osteoregenerative agents. OST gene sequence can be used
CC for assessing the osteoregenerative activity of a test agent in a
CC subject, and for diagnosing or determining the susceptibility to bone
CC disorder and assessing efficacy of a treatment of a bone disorder in a
CC subject, human or rodent. (M2) is useful for treating a bone disorder
CC including bone formation disorder or bone resorption disorder (e.g.
CC osteoporosis, osteopenia and osteopetrosis). OST polypeptides are useful
CC as immunogens to raise anti-OST antibodies (II). Parathyroid hormone
CC receptor ligands (I) identified by (M1) are useful for treating bone
CC disorders. The present sequence represents an OST23 gene fragment
CC isolated from rat tibiae, from the present invention.
XX
SQ Sequence 439 BP; 128 A; 108 C; 99 G; 104 T; 0 other;

Query Match 69.1%; Score 121; DB 24; Length 439;
Best Local Similarity 100.0%; Pred. No. 8.8e-52;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCAGGAAGCTTGTGATTTTTCAGGAATGACATCTCTTAAGCACTCGCAAAAC 60
Db 285 AGTCAGGAAGCTTGTGATTTTTCAGGAATGACATCTCTTAAGCACTCGCAAAAC 344

QY 61 AGGAAGGCTCCACACCTCTGCGAGGCGGCTTCTCTTCAGCATGAGAAAGCAAGG 120
Db 345 AGGAAGGCTCCACACCTCTGCGAGGCGGCTTCTCTTCAGCATGAGAAAGCAAGG 404

QY 121 G 121
Db 405 G 405

RESULT 3
ABN30189
ID ABN30189 standard; DNA; 65 BP.
XX
AC ABN30189;
XX
DT 15-JUL-2002 (first entry)
XX
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:2937.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Rattus norvegicus.
XX
PN WO200210449-A2.
PN
PD 07-FEB-2002.
PD
XX
PF 20-JUL-2001; 2001WO-IB01903.
PF
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
PR
XX
PA (COMP-) COMPUGEN INC.
PA
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI
XX
DR WPI; 2002-257383/30.
DR
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes

```
XX PS Example 1; SEQ ID 2937; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 21 A; 20 C; 13 G; 11 T; 0 other;

Query Match
Best Local Similarity 30.9%; Score 54; DB 24; Length 65;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 TTCAGGATGCACATCTCTTAAGCACTCGCAAAACAGGAGGCTCCACACCTCT 79
DB 1 TTCAGGATGCACATCTCTTAAGCACTCGCAAAACAGGAGGCTCCACACCTCT 54

RESULT 4
AAAX76488
ID AAX76488 standard; DNA; 1734 BP.
XX
AC AAX76488;
XX
DT 06-AUG-1999 (first entry)
XX
DE Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Mus sp.
XX
PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
```

```
XX DR WPI; 1999-337420/28.
XX DR P-PSDB; AAY17651.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Example 2; Page 178-179; 284pp; English.
XX
CC The present invention describes Wnt-1 induced secreted polypeptides.
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;

Query Match
Best Local Similarity 20.0%; Score 35; DB 20; Length 1734;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGA 111
DB 1611 TCTGGCAGGCCAGGGCCTTCTCTTCAGCATGAGA 1645

RESULT 5
AAD32413/C
ID AAD32413 standard; DNA; 1668 BP.
XX
AC AAD32413;
XX
DT 18-JUN-2002 (first entry)
XX
DE Chlamydia pneumoniae AR039 omcB/ompB gene.
XX
KW Chlamydiaceae family; chronic infection; persistent infection; pyk; nlpD;
KW Cpn0585; regulatory pathway; biosynthetic pathway; ompA; ompB; hsp60;
KW lipopolysaccharide; cardiovascular system; respiratory tract; therapy;
KW genital tract; reproductive system; atherosclerotic tissue; macrophage;
KW multiple sclerosis; conjunctiva; prophylaxis; antibacterial; gene; ds.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FH CDS 1..1168
FT /*tag= a
FT /product= "Chlamydia pneumoniae AR039 omcB/ompB protein"
FT /note= "CDS does not include stop codon"
FT /partial
XX
PN WO200214516-A1.
XX
PD 21-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-AU01021.
XX
PR 18-AUG-2000; 2000AU-0009540.
XX
```

PA (UYOU-) UNIV QUEENSLAND TECHNOLOGY.
PA (MATH/) MATHEWS S A.
XX
PI Tlmm P;
XX
DR WPI: 2002-269197/31.
DR P-PSDB; AAE20302.
XX
PT Detecting Chlamydial organism in its persistent phase by detecting
PT expression change of range of genes belonging to their respective
PT biosynthetic pathways when expression is compared to that of organism
PT in lytic phase -
XX
PS Disclosure: Page 168-170; 196pp; English.
XX
CC The invention relates to composition and methods for detecting organisms
CC of the Chlamydiaceae family, including species of Chlamydia and
CC Chlamydia, in the persistent phase of their developmental cycle and
CC for the diagnosis of chronic or persistent infections caused by such
CC organisms. The composition is useful for modulating the expression of
CC gene such as pyk, nlpD, Cpn0585, a gene belonging to same regulatory/
CC biosynthetic pathway and ompA, ompB, hsp60, a gene involved in
CC lipopolysaccharide biosynthesis. It is also useful for modulating the
CC the level and/or functional activity of an expression product of these
CC genes, where the gene is present in an epithelial cell (selected from
CC cardiovascular system, respiratory tract, genital tract, reproductive
CC system or conjunctiva), macrophage, or a cell associated with
CC atherosclerotic tissue or associated with multiple sclerosis brain
CC tissue. The composition is useful for treatment and/or prophylaxis of a
CC chronic infection caused by an organism of the Chlamydiaceae family in
CC a patient. Antigen associated with the persistent phase of the
CC developmental cycle of an organism of the Chlamydiaceae family, is
CC useful in the manufacture of a medicament, for treating and/or
CC preventing Chlamydia infection in a patient. The present sequence
CC is Chlamydia pneumoniae AR039 omcB/ompB gene.
XX
SQ Sequence 1668 BP; 486 A; 342 C; 370 G; 470 T; 0 other;
XX
Query Match 10.3%; Score 18; DB 24; Length 1668;
Best Local Similarity 100.0%; Pred. NO. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 TCCAGGAACTTGAGCTTT 20
Db 1227 TCCAGGAACTTGAGCTTT 1210
Db 1227 TCCAGGAACTTGAGCTTT 1210
RESULT 6
ABL91207/C
ID ABL91207 standard; DNA; 1671 BP.
XX
AC ABL91207;
XX
DT 29-JUL-2002 (first entry)
XX
DE Chlamydia pneumoniae cp6849 ORF DNA, SEQ ID NO:48.
XX
KW Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
KW human respiratory disease; cardiovascular disease; atherosclerosis;
KW coronary artery disease; carotid artery stenosis; myocardial infarction;
KW cerebrovascular disease; aortic aneurysm; claudication; stroke;
KW strain CWL029; open reading frame; ORF; gene; ds.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 1..1671
FT /*tag- a
FT /*product= "cp6849"
FT sig_peptide 1..66
FT /*tag- b
FT mat_peptide 67..1668
FT /*tag- c
XX

FT /*product= "Mature protein"
XX
PN WO200202606-A2.
XX
PD 10-JAN-2002.
XX
PE 03-JUL-2001; 2001WO-IB01445.
XX
PR 03-JUL-2000; 2000GB-0016363.
PR 11-JUL-2000; 2000GB-0017047.
PR 21-JUL-2000; 2000GB-0017983.
PR 07-AUG-2000; 2000GB-0019368.
PR 18-AUG-2000; 2000GB-0020440.
PR 14-SEP-2000; 2000GB-0022583.
PR 10-NOV-2000; 2000GB-0027549.
PR 22-DEC-2000; 2000GB-0031706.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Ratti G, Grandi G;
XX
XX WPI: 2002-154726/20.
DR N-PSDB; ABB90549.
XX
PT Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT medicament for treatment or prevention of infection due to Chlamydia,
PT preferably Chlamydia pneumoniae, and for diagnostic purposes -
XX
PS Claim 5; Page 66; 364pp; English.
XX
CC Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
CC pneumoniae (strain CWL029), and ABL91184-ABL91373 represent DNA encoding
CC them. The proteins are predicted to be immunogenic and may therefore be
CC useful in vaccine production and for diagnostic purposes. Chlamydia
CC pneumoniae is a common cause of respiratory disease in humans, and is
CC also involved in the development of cardiovascular diseases such as
CC atherosclerosis, coronary artery disease, carotid artery stenosis,
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC claudication and stroke. The proteins and nucleic acids of the invention
CC may be used in vaccines and pharmaceutical compositions for the
CC prevention or treatment of chlamydial infections, particularly Chlamydia
CC pneumoniae infections. The proteins may also be used in the detection of
CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
CC DNA probe assay or blotting techniques for determining Chlamydia
CC pneumoniae gene expression. The present sequence represents a
CC specifically claimed DNA which encodes a Chlamydia pneumoniae protein of
CC the invention.
XX
SQ Sequence 1671 BP; 488 A; 342 C; 370 G; 471 T; 0 other;
XX
Query Match 10.3%; Score 18; DB 24; Length 1671;
Best Local Similarity 100.0%; Pred. NO. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 TCCAGGAACTTGAGCTTT 20
Db 1227 TCCAGGAACTTGAGCTTT 1210
Db 1227 TCCAGGAACTTGAGCTTT 1210
RESULT 7
ABL92605/C
ID ABL92605 standard; DNA; 1671 BP.
XX
AC ABL92605;
XX
DT 05-JUN-2002 (first entry)
XX
DE Chlamydia pneumoniae DNA sequence SEQ ID NO:378.
XX
KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis;
KW antigen; antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response; gene; ds.
XX

OS	Chlamydia pneumoniae.
XX	
PN	WO200208267-A2.
XX	
PD	31-JAN-2002.
XX	
PF	20-JUL-2001; 2001WO-US23121.
XX	
PR	20-JUL-2000; 2000US-0620412.
PR	23-APR-2001; 2001US-0841132.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Fling SP, Skelky YAW, Probst P, Bhatia A;
XX	
DR	WPI; 2002-179901/23.
XX	
PT	Novel compositions comprising Chlamydia Cap1 protein and its use in the
PT	treatment of Chlamydia infection -
XX	
PS	Disclosure; Page 333; 537pp; English.
XX	
CC	The present invention describes compositions comprising a Chlamydia Cap1
CC	protein and methods for the diagnosis and therapy of Chlamydia infection.
CC	Chlamydia DNA and protein sequences from the present invention can have
CC	antibacterial and immunostimulant activities, and can be used in
CC	vaccines. Compounds from the present invention can be used for eliciting
CC	an immune response, specifically stimulating a Chlamydia-specific T-cell
CC	response or inhibiting the development of a Chlamydia infection in an
CC	animal. Methods from the present invention can be used: for detecting the
CC	presence of Chlamydia in a patient; to stimulate and/or expand T cells
CC	specific for a Chlamydia protein; and for treatment of a Chlamydia
CC	infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
CC	sequences used in the exemplification of the present invention.
XX	
SO	Sequence 1671 BP; 488 A; 342 C; 370 G; 471 T; 0 other;
	Query Match 10.3%; Score 18; DB 24; Length 1671;
	Best Local Similarity 100.0%; Pred. No. 23;
	Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	3 TCCAGGAAGCTGAGCTTT 20
Db	1227 TCCAGGAAGCTGAGCTTT 1210
	RESULT 8
	ABL91770/c
ID	ABL91770 standard; DNA; 2031 BP.
XX	
AC	ABL91770;
XX	
DT	28-MAY-2002 (first entry)
XX	
DE	Human polynucleotide SEQ ID NO 113.
XX	
KW	Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW	Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW	cytostatic; virucide; protozoacide; antibacterial; ds.
XX	
OS	Homo sapiens.
XX	
PN	DE10100586-C1.
XX	
PD	11-APR-2002.
XX	
PF	09-JAN-2001; 2001DE-1000586.
XX	
PR	09-JAN-2001; 2001DE-1000586.
XX	
PA	(RIBO-) RIBOPHARMA AG.
XX	
PI	Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX	WPI; 2002-270454/32.
XX	
PT	Inhibiting gene expression in cells, useful for e.g. treating tumors,
PT	by introducing double-stranded complementary oligoRNA having unpaired
PT	terminal bases -
XX	
PS	Claim 13; Page 83-84; 104pp; German.
XX	
CC	The invention relates to a method for inhibiting expression of a target
CC	gene (AB191658-AB191797) in a cell by introducing at least one
CC	oligonucleotide that has a double-stranded structure consisting of at
CC	most 49 sequential nucleotide pairs, with at least part of one strand
CC	complementary with the target gene and has at least one end a
CC	single-stranded segment of 1-4 nt. The method provides
CC	oligoribonucleotides for antisense inhibition of gene expression useful
CC	e.g. for treating tumours but the oligoribonucleotides may also be
CC	directed against genes present in pathogens (e.g. Plasmodium or
CC	viruses/viroids, pathogenic on humans, animals or plants), or against
CC	cytokine, id, developmental or prion genes. The method provides more
CC	effective inhibition of gene expression than use of known
CC	oligonucleotides, probably because the unpaired overhang increases
CC	stability and thus intracellular concentration.
XX	
SQ	Sequence 2031 BP; 514 A; 549 C; 561 G; 407 T; 0 other;
	Query Match 10.3%; Score 18; DB 24; Length 2031;
	Best Local Similarity 100.0%; Pred. No. 23;
	Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	125 GCAGAGTACTCTCTCTG 142
Db	1127 GCAGAGTACTCTCTCTG 1110
	RESULT 9
ID	AAF21412/c
XX	AAF21412 standard; DNA; 2104 BP.
AC	AAF21412;
XX	
DT	14-MAR-2001 (first entry)
XX	
DE	Human low adenosine antisense oligonucleotide related sequence #2979.
XX	
KW	Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW	human; airway disorder; bronchoconstriction; lung inflammation;
KW	surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW	immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW	respiratory obstruction; pulmonary obstruction; impeded respiration;
KW	surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW	respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW	pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW	chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW	cancer; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200062736-A2.
XX	
PD	26-OCT-2000.
XX	
PF	24-MAR-2000; 2000WO-US08020.
XX	
PR	06-APR-1999; 99US-0127958.
XX	
PA	(UYEC-) UNIV EAST CAROLINA.
PA	(NYCE/) NYCE J W.
PI	
PI	Nyce JW;
DR	WPI; 2000-679539/66.
XX	

PT Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
XX
PS Disclosure; Page 1393; 1592pp; English.

CC The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, anti-inflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.

XX
SQ Sequence 2104 BP; 524 A; 581 C; 582 G; 417 T; 0 other;

Query Match 10.3%; Score 18; DB 21; Length 2104;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 GCAGAGTACTCTCTCTG 142
|||||
DB 1185 GCAGAGTACTCTCTCTG 1168

RESULT 10
AAA35290/C
ID AAA35290 standard; DNA; 2104 BP.
XX
AC AAA35290;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:164.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impeded respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; anti-inflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17712.

XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -
XX
XX
PS Disclosure; Page 1308; 1343pp; English.

CC The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have anti-inflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.

XX
SQ Sequence 2104 BP; 524 A; 581 C; 582 G; 417 T; 0 other;

Query Match 10.3%; Score 18; DB 21; Length 2104;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 GCAGAGTACTCTCTCTG 142
|||||
DB 1185 GCAGAGTACTCTCTCTG 1168

RESULT 11
AAC84222/C
ID AAC84222 standard; CDNA; 2104 BP.
XX
AC AAC84222;
XX
DT 19-MAR-2001 (first entry)
XX
DE Human protein kinase C-delta (PKC-delta) encoding CDNA.
XX
KW Antisense; human; protein kinase C-delta; PKC-delta; antipsoriatic;
KW cytostatic; immunosuppressive; antidiabetic; neuroprotective; cancer;
KW anti-inflammatory; antirheumatic; antiarthritic; hepatotropic; hepatitis;
KW inflammatory bowel disease; multiple sclerosis; pancreatitis; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
59..2089
/*tag= a

```
FT XX /product= "PKC-delta"
XX PN WO200070091-A1.
XX PD 23-NOV-2000.
XX PF 11-MAY-2000; 2000WO-US13170.
XX PR 18-MAY-1999; 99US-0313930.
XX PA (ISIS-) ISIS PHARM INC.
XX PI Dean NM;
XX DR WPI; 2001-025027/03.
DR P-PSDB; AAB48037.
XX
PT Antisense oligonucleotide for modulating human protein kinase C-delta
PT and treating diseases such as cancer, diabetes, comprises nucleotides
PT specifically hybridizable with nucleic acid encoding the protein kinase
PT
XX
PS Example 2; Page 70-74; 84pp; English.
XX
CC The invention provides antisense oligonucleotides that specifically
CC hybridize with human protein kinase C-delta (PKC-delta) polynucleotide.
CC The antisense oligonucleotides are useful for modulating the expression
CC of human PKC-delta in cells or tissues, especially adipose tissues which
CC express human PKC-delta, for inhibiting hyperproliferation of cells and
CC treating or preventing an abnormal proliferative condition such as
CC hyperproliferative disorder, especially psoriasis or cancer such as
CC leukemia. They are also useful for reducing an inflammatory response of
CC human cells and treating an animal having a disease or condition
CC associated with tumour necrosis factor-alpha (TNF-alpha), especially
CC overexpression of TNF-alpha, such as inflammatory or autoimmune disease
CC or a condition which is infectious, diabetes, inflammatory bowel
CC disease, multiple sclerosis, pancreatitis, rheumatoid arthritis,
CC hepatitis, atopic dermatitis or allograft rejection and for reducing
CC the blood glucose level in a human. The present sequence represents the
CC cDNA encoding the human PKC-delta polypeptide.
XX
SQ Sequence 2104 BP; 524 A; 581 C; 582 G; 417 T; 0 other;

Query Match 10.3%; Score 18; DB 22; Length 2104;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 GCAGAGTACTCTCCTCTG 142
Db 1185 GCAGAGTACTCTCCTCTG 1168

RESULT 12
AAA75901/C
ID AAA75901 standard; DNA; 2111 BP.
XX
AC AAA75901;
XX
DT 22-JAN-2001 (first entry)
XX
DE DNA encoding a 60 kDa cysteine-rich membrane protein.
XX
KW Cysteine-rich membrane protein; Chlamydia infection; bronchitis;
KW community acquired pneumonia; upper respiratory tract infection; vaccine;
KW sinusitis; ss.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 139..1809
FT /*tag= a
FT /product= "membrane-rich protein"
XX
```

```
PN WO200055326-A1.
XX
XX PD 21-SEP-2000.
XX PF 09-MAR-2000; 2000WO-CA00240.
XX PR 12-MAR-1999; 99US-0123966.
XX PA (AVET ) AVENTIS PASTEUR LTD.
XX PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX DR WPI; 2000-618918/59.
DR P-PSDB; AAB18820.
XX
PT New polynucleotides encoding a 60kda cysteine-rich membrane protein
PT from Chlamydia, useful as a vaccine for preventing and treating
PT Chlamydia infection in mammals
XX
PS Claim 2; Fig 1; 77pp; English.
XX
CC The present sequence encodes a Chlamydia 60 kDa cysteine-rich membrane
CC protein. The membrane-rich polynucleotide and polypeptide are useful
CC for preventing or treating Chlamydia infection, such as community
CC pneumonia, upper respiratory tract infections, bronchitis and sinusitis.
CC They are also useful for diagnosing Chlamydia infection by assaying a
CC body fluid of a mammal. The polypeptide is useful for vaccine
CC production.
XX
SQ Sequence 2111 BP; 634 A; 417 C; 442 G; 618 T; 0 other;

Query Match 10.3%; Score 18; DB 21; Length 2111;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TCCAGGAACTTGAGCTTT 20
Db 1365 TCCAGGAACTTGAGCTTT 1348

RESULT 13
AAS56353
ID AAS56353 standard; DNA; 2586 BP.
XX
AC AAS56353;
XX
DT 13-FEB-2002 (first entry)
XX
DE Salmonella typhi DNA for cellular proliferation protein #386.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Salmonella typhi.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
FH Haselbeck R, Ohlsen KL, zyskind JW, Wall D, Trawick JD, Carr GJ;
FT Yamamoto RT, Xu HH;
PI
```

XX WPI; 2001-611495/70.
 DR P-PSDB; AAU38494.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Claim 27; Seq ID No 9990; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 2586 BP; 642 A; 618 C; 764 G; 561 T; 1 other;
 Query Match 10.3%; Score 18; DB 23; Length 2586;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 CTCGCAAAACAGCAAGC 68
 |||||||
 Db 808 CTCGCAAAACAGCAAGC 825

RESULT 14
 AAH57492
 ID AAH57492 standard; cDNA; 5095 BP.
 XX
 AC AAH57492;
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Human liver cell specific cDNA sequence SEQ ID NO:332.
 XX
 KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
 KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
 KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
 KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
 XX
 OS Homo sapiens.
 XX
 PN WO200132927-A2.
 XX
 PD 10-MAY-2001.
 XX
 PF 02-NOV-2000; 2000WO-US30396.
 XX
 PR 04-NOV-1999; 99US-0163508.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Sornasse T, Selthamer JJ, Watson GA;
 XX
 DR WPI; 2001-291057/30.
 XX
 PT New cell and tissue specific polynucleotides useful for diagnosis,
 PT prognosis or monitoring of treatments for disorders where the gene is

PT associated with a cancer, immunopathology or neuropathology -
 XX
 PS Claim 1; Page 251-252; 327pp; English.
 XX
 CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
 CC sequences (I). (I) can have cytostatic, immunomodulatory and
 CC neuroprotective activities, and can be used in gene therapy. (I) and
 CC proteins (II) encoded by then are used in high throughput screening
 CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
 CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
 CC their fragments, immunoglobulins, inhibitors, drug compounds and
 CC pharmaceutical agents. Expression of (I) in a sample indicates the
 CC differentiation of embryonic stem cells into a tissue selected from
 CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
 CC tissues. (I) and (II) are used to produce an expression profile that
 CC defines a metabolic or developmental process, treatment, condition,
 CC disease or disorder. The gene profile can be used for diagnosis,
 CC prognosis or monitoring of treatments and for investigating a
 CC predisposition to a disorder where the gene is associated with a
 CC cancer, immunopathology or neuropathology.
 XX
 SO Sequence 5095 BP; 1596 A; 867 C; 941 G; 1691 T; 0 other;
 Query Match 10.3%; Score 18; DB 22; Length 5095;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 74 ACCTCTGGCAGCGCAGG 91
 |||||||
 Db 3636 ACCTCTGGCAGCGCAGG 3653

RESULT 15
 AAF21424/C
 ID AAF21424 standard; DNA; 38644 BP.
 XX
 AC AAF21424;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Human low adenosine antisense oligonucleotide related sequence #2991.
 XX
 KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200062736-A2.
 XX
 PD 26-OCT-2000.
 XX
 PF 24-MAR-2000; 2000WO-US08020.
 XX
 PR 06-APR-1999; 99US-0127958.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 PA (NYCE/) NYCE J W.
 XX
 PI Nyce JW;
 XX
 DR WPI; 2000-679539/66.
 XX
 PT Low adenosine (A) content antisense oligonucleotides which do not
 PT trigger adenosine receptors during metabolism, useful e.g. for treating
 PT cancers and respiratory obstructions -

XX
PS Disclosure; Page 1400-1409; 1592pp; English.

CC The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.

XX
SQ Sequence 38644 BP; 9551 A; 10265 C; 10290 G; 8538 T; 0 other;

Query Match 10.3%; Score 18; DB 21; Length 38644;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 GCAGAGTACTCTCTCTG 142
|||||
DB 16814 GCAGAGTACTCTCTCTG 16797

Search completed: July 28, 2003, 15:54:59
Job time : 47.9617 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:50:06 ; Search time 9.56307 Seconds
(without alignments)
5612.050 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708

Perfect score: 175
Sequence: 1 AGTCAGAGACTTGAGCTTT.....GCCTAGATTAACACCCCAA 175

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	20.0	1734	4	US-09-182-145-17
2	35	20.0	1734	4	US-09-182-145-18
3	18	10.3	2104	4	US-09-313-930-1
4	17	9.7	24	4	US-09-182-145-110
5	17	9.7	742	2	US-08-966-316-2
6	16	9.1	463	4	US-09-556-877-55
7	16	9.1	463	4	US-09-556-877-58
8	16	9.1	463	4	US-09-620-412C-55
9	16	9.1	463	4	US-09-620-412C-58
10	16	9.1	601	4	US-09-556-877-22
11	16	9.1	601	4	US-09-288-594A-22
12	16	9.1	601	4	US-09-620-412C-22
13	16	9.1	696	4	US-09-556-877-25
14	16	9.1	696	4	US-09-288-594A-25
15	16	9.1	696	4	US-09-620-412C-25
16	16	9.1	1256	4	US-09-556-877-21
17	16	9.1	1256	4	US-09-288-594A-21
18	16	9.1	1256	4	US-09-620-412C-21
19	16	9.1	2192	3	US-08-942-001-1
20	16	9.1	2192	4	US-09-337-386-1
21	16	9.1	2192	4	US-09-846-922-1
22	15	8.6	282	4	US-09-375-318-15
23	15	8.6	737	4	US-08-858-207A-183
24	15	8.6	1536	4	US-09-268-992-44
25	15	8.6	1536	4	US-09-657-474-44
26	15	8.6	1539	4	US-09-268-992-42
27	15	8.6	1539	4	US-09-657-474-42

C	28	15	8.6	1767	4	US-09-268-992-40	Sequence 40, Appl
C	29	15	8.6	1767	4	US-09-657-474-40	Sequence 40, Appl
C	30	15	8.6	1815	4	US-09-268-992-38	Sequence 38, Appl
C	31	15	8.6	1815	4	US-09-657-474-38	Sequence 38, Appl
	32	15	8.6	2404	1	US-08-484-105-7	Sequence 7, Appl1
	33	15	8.6	2404	1	US-08-484-106-7	Sequence 7, Appl1
C	34	15	8.6	2483	1	US-08-464-340A-3	Sequence 3, Appl1
C	35	15	8.6	2483	5	PCT-US94-08449A-3	Sequence 3, Appl1
C	36	15	8.6	2730	2	US-08-811-897A-39	Sequence 39, Appl
C	37	15	8.6	2730	2	US-08-855-213-39	Sequence 39, Appl
C	38	15	8.6	2730	4	US-09-201-474-39	Sequence 39, Appl
C	39	15	8.6	2814	2	US-08-811-897A-38	Sequence 38, Appl
C	40	15	8.6	2814	2	US-08-855-213-38	Sequence 38, Appl
C	41	15	8.6	2814	4	US-09-201-474-38	Sequence 38, Appl
C	42	15	8.6	2940	4	US-09-703-416-1	Sequence 1, Appl1
C	43	15	8.6	3748	2	US-08-958-240-1	Sequence 1, Appl1
C	44	15	8.6	4188	4	US-07-751-891B-2	Sequence 2, Appl1
	45	15	8.6	4242	3	US-07-705-490-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-182-145-17
; Sequence 17, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-182-145-17

Query Match      20.0%; Score 35; DB 4; Length 1734;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      77 TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA 111
Db      1611 TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA 1645

RESULT 2
US-09-182-145-18/c
; Sequence 18, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
```



```

; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/966,316
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0424 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 742 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLNFET02
; CLONE: 1457779
; US-08-966-316-2

```

```

Query Match          9.7%; Score 17; DB 2; Length 742;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      78 CTGGCAGGCCAGGGCCT 94
      ||||||||||||||||
DB      428 CTGGCAGGCCAGGGCCT 412

```

```

RESULT 6
US-09-556-877-55/c
; Sequence 55, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 55
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
; US-09-556-877-55

```

```

Query Match          9.1%; Score 16; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      117 AAGGACAGCAGAGTA 132
      ||||||||||||||||
DB      412 AAGGACAGCAGAGTA 397

```

```

RESULT 7
US-09-556-877-58/c
; Sequence 58, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 58
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
; US-09-556-877-58

```

```

Query Match          9.1%; Score 16; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      117 AAGGACAGCAGAGTA 132
      ||||||||||||||||
DB      412 AAGGACAGCAGAGTA 397

```

```

RESULT 8
US-09-620-412C-55/c
; Sequence 55, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7.
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 55
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
; US-09-620-412C-55

```

```

Query Match          9.1%; Score 16; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      117 AAGGACAGCAGAGTA 132
      ||||||||||||||||
DB      412 AAGGACAGCAGAGTA 397

```

```

RESULT 9
US-09-620-412C-58/c
; Sequence 58, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0

```



```
; SEQ ID NO 58
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-58
```

```
Query Match          9.1%; Score 16; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      117 AAGGACAGCAGAGTA 132
          |||
Db       412 AAGGACAGCAGAGTA 397
```

RESULT 10

```
US-09-556-877-22/c
; Sequence 22, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 22
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-556-877-22
```

```
Query Match          9.1%; Score 16; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      117 AAGGACAGCAGAGTA 132
          |||
Db       537 AAGGACAGCAGAGTA 522
```

RESULT 11

```
US-09-288-594A-22/c
; Sequence 22, Application US/09288594A
; Patent No. 6447779
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C1
; CURRENT APPLICATION NUMBER: US/09/288,594A
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-288-594A-22
```

```
Query Match          9.1%; Score 16; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      117 AAGGACAGCAGAGTA 132
          |||
Db       537 AAGGACAGCAGAGTA 522
```

RESULT 12

```
US-09-620-412C-22/c
; Sequence 22, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 22
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-22
```

```
Query Match          9.1%; Score 16; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      117 AAGGACAGCAGAGTA 132
          |||
Db       537 AAGGACAGCAGAGTA 522
```

RESULT 13

```
US-09-556-877-25/c
; Sequence 25, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 25
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-556-877-25
```

```
Query Match          9.1%; Score 16; DB 4; Length 696;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      117 AAGGACAGCAGAGTA 132
          |||
Db       538 AAGGACAGCAGAGTA 523
```

RESULT 14

```
US-09-288-594A-25/c
; Sequence 25, Application US/09288594A
; Patent No. 6447779
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
```

APPLICANT: Fling, Steven P.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE TREATMENT
TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 210121.469C1
CURRENT APPLICATION NUMBER: US/09/288,594A
CURRENT FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 25
LENGTH: 696
TYPE: DNA
ORGANISM: Chlamydia trachomatis
US-09-288-594A-25

Query Match 9.1%; Score 16; DB 4; Length 696;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 AAGGACAGCAGAGTA 132
|||||
DB 538 AAGGACAGCAGAGTA 523

RESULT 15
US-09-620-412C-25/C
Sequence 25, Application US/09620412C
Patent No. 6448234
GENERAL INFORMATION:
APPLICANT: Steven P. Fling
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 210121.469C7
CURRENT APPLICATION NUMBER: US/09/620,412C
CURRENT FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 25
LENGTH: 696
TYPE: DNA
ORGANISM: Chlamydia trachomatis
US-09-620-412C-25

Query Match 9.1%; Score 16; DB 4; Length 696;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 AAGGACAGCAGAGTA 132
|||||
DB 538 AAGGACAGCAGAGTA 523

Search completed: July 28, 2003, 15:58:55
Job time : 10.5631 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:30:21 ; Search time 43.097 Seconds
(without alignments)
8377.033 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708
Perfect score: 175
Sequence: 1 AGTCCAGGAAGCTTGAGCTTT.....GCCTAGAAATAACACCCAAA 175

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1439767 seqs, 1031500376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	175	100.0	1708	15	US-10-010-408-1
2	121	69.1	439	12	US-09-956-622A-23
3	35	20.0	1734	15	US-10-112-267-17
4	35	20.0	1734	15	US-10-112-267-18
5	18	10.3	1671	10	US-09-841-132-378
6	18	10.3	2586	10	US-09-815-242-9990
7	18	10.3	3073	15	US-10-027-632-114289
8	18	10.3	3073	15	US-10-027-632-114290
9	18	10.3	3186778	15	US-10-027-632-174961
10	17	9.7	24	15	US-10-112-267-110
11	17	9.7	141	11	US-09-783-590-7073
12	17	9.7	451	12	US-09-918-995-12392
13	17	9.7	627	15	US-10-027-632-204519
14	17	9.7	627	15	US-10-027-632-204520
15	17	9.7	627	15	US-10-027-632-204521
16	17	9.7	738	15	US-10-027-632-32541

C	17	17	9.7	742	12	US-09-968-433-2	Sequence 2, Appli
	18	17	9.7	774	15	US-10-027-632-114990	Sequence 114990,
	19	17	9.7	774	15	US-10-027-632-114991	Sequence 114991,
	20	17	9.7	940	15	US-10-027-632-261119	Sequence 261119,
	21	17	9.7	940	15	US-10-027-632-261120	Sequence 261120,
	22	17	9.7	2126	15	US-10-052-586-495	Sequence 495, App
	23	17	9.7	2126	15	US-10-174-590-495	Sequence 495, App
	24	17	9.7	2126	15	US-10-176-758-495	Sequence 495, App
	25	17	9.7	2126	15	US-10-175-737-495	Sequence 495, App
	26	17	9.7	2126	15	US-10-173-706-495	Sequence 495, App
	27	17	9.7	2126	15	US-10-175-738-495	Sequence 495, App
	28	17	9.7	2126	15	US-10-175-752-495	Sequence 495, App
	29	17	9.7	2126	15	US-10-176-482-495	Sequence 495, App
	30	17	9.7	2126	15	US-10-176-757-495	Sequence 495, App
	31	17	9.7	2126	15	US-10-176-913-495	Sequence 495, App
	32	17	9.7	2126	15	US-10-180-552-495	Sequence 495, App
	33	17	9.7	2126	15	US-10-180-557-495	Sequence 495, App
	34	17	9.7	2126	15	US-10-173-700-495	Sequence 495, App
	35	17	9.7	2126	15	US-10-174-572-495	Sequence 495, App
	36	17	9.7	2126	15	US-10-174-579-495	Sequence 495, App
	37	17	9.7	2126	15	US-10-174-582-495	Sequence 495, App
	38	17	9.7	2126	15	US-10-174-588-495	Sequence 495, App
	39	17	9.7	2126	15	US-10-175-739-495	Sequence 495, App
	40	17	9.7	2126	15	US-10-175-740-495	Sequence 495, App
	41	17	9.7	2126	15	US-10-175-743-495	Sequence 495, App
	42	17	9.7	2126	15	US-10-176-488-495	Sequence 495, App
	43	17	9.7	2126	15	US-10-176-492-495	Sequence 495, App
	44	17	9.7	2126	15	US-10-176-747-495	Sequence 495, App
	45	17	9.7	2126	15	US-10-176-750-495	Sequence 495, App

ALIGNMENTS

RESULT 1

US-10-010-408-1
Sequence 1, Application US/10010408
Publication No. US20020165185A1

GENERAL INFORMATION:

APPLICANT: John J. Castellot, Jr.
TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules and Uses Therefor

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston

STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 1708 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 249..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-010-408-1

Query Match 100.0%; Score 175; DB 15; Length 1708;
Best Local Similarity 100.0%; Pred. No. 3.1e-85;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGTCAGGAACCTTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
Db 1534 AGTCAGGAACCTTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 1593
OY 61 AGGAAGGCTCCACACCTCTGGCAGGCGCCCTTCTCTTCAGCATGAGAAAGACAAGG 120
Db 1594 AGGAAGGCTCCACACCTCTGGCAGGCGCCCTTCTCTTCAGCATGAGAAAGACAAGG 1653
OY 121 GACAGCAGAGTACTCTCTCTGGAGGAGTACTAGTCTAGATTAACACCCCAA 175
Db 1654 GACAGCAGAGTACTCTCTCTGGAGGAGTACTAGTCTAGATTAACACCCCAA 1708

RESULT 2

US-09-956-622A-23
; Sequence 23, Application US/09956622A
; Publication No. US20030091973A1
; GENERAL INFORMATION:
; APPLICANT: Horesovsky, Gregory J
; APPLICANT: No. US20030091973A1 II, L. Staton
; APPLICANT: Raha, Debashish
; TITLE OF INVENTION: Method of Identifying Osteoregenerative Agents Using
; FILE REFERENCE: 21402-445
; CURRENT APPLICATION NUMBER: US/09/956,622A
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/233,579
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-956-622A-23

Query Match 69.1%; Score 121; DB 12; Length 439;
Best Local Similarity 100.0%; Pred. No. 7.4e-56;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGTCAGGAACCTTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
Db 285 AGTCAGGAACCTTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 344
OY 61 AGGAAGGCTCCACACCTCTGGCAGGCGCCCTTCTCTTCAGCATGAGAAAGACAAGG 120
Db 345 AGGAAGGCTCCACACCTCTGGCAGGCGCCCTTCTCTTCAGCATGAGAAAGACAAGG 404
OY 121 G 121
Db 405 G 405

RESULT 3

US-10-112-267-17
; Sequence 17, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:

APPLICANT: Botstein, David A.
APPLICANT: Cohen, Robert
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Lawrence, David A.
APPLICANT: Levine, Arnold J.
APPLICANT: Pennica, Diane
APPLICANT: Roy, Margaret Ann
APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-112-267-17

Query Match 20.0%; Score 35; DB 15; Length 1734;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 TCTGGCAGGCGCCGCTTCTCTTCAGCATGAGA 111
Db 1611 TCTGGCAGGCGCCGCTTCTCTTCAGCATGAGA 1645

RESULT 4

US-10-112-267-18/c
; Sequence 18, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 18
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-112-267-18

Query Match 20.0%; Score 35; DB 15; Length 1734;

Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 77 TCTGGCAGCCAGGGCCTTCTCTTCAGCATGAGA 111
DB 124 TCTGGCAGCCAGGGCCTTCTCTTCAGCATGAGA 90

RESULT 5

US-09-841-132-378/c
; Sequence 378, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 378
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-378

Query Match 10.3%; Score 18; DB 10; Length 1671;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TCCAGGAAGCTTGAGCTTT 20
DB 1227 TCCAGGAAGCTTGAGCTTT 1210

RESULT 6

US-09-815-242-9990
; Sequence 9990, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9990

; LENGTH: 2586
; TYPE: DNA
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2586)
; NAME/KEY: misc_feature
; LOCATION: (1)...(2586)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-242-9990

Query Match 10.3%; Score 18; DB 10; Length 2586;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 CTCGCAAAACAGGAAGGC 68
DB 808 CTCGCAAAACAGGAAGGC 825

RESULT 7

US-10-027-632-114289
; Sequence 114289, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114289
; LENGTH: 3073
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-114289

Query Match 10.3%; Score 18; DB 15; Length 3073;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 AAGGCTCCACACCTCTGG 81
DB 2440 AAGGCTCCACACCTCTGG 2457

RESULT 8

US-10-027-632-114290
; Sequence 114290, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 114290
LENGTH: 3073
TYPE: DNA
ORGANISM: Human
US-10-027-632-114290

Query Match 10.3%; Score 18; DB 15; Length 3073;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 AAGCTCCACACCTCTGG 81
|||||
DB 2440 AAGCTCCACACCTCTGG 2457

RESULT 9
US-10-027-632-174961

Sequence 174961, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 174961
LENGTH: 3186778
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(3186778)
OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174961

Query Match 10.3%; Score 18; DB 15; Length 3186778;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 AGCACTCGCAACAGCA 64
|||||
DB 805006 AGCACTCGCAACAGCA 805023

RESULT 10

US-10-112-267-110/c
Sequence 110, Application US/10112267
Publication No. US20030068678A1
GENERAL INFORMATION:
APPLICANT: Bolstein, David A.
APPLICANT: Cohen, Robert
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Lawrence, David A.
APPLICANT: Levine, Arnold J.
APPLICANT: Pennica, Diane
APPLICANT: Roy, Margaret Ann
APPLICANT: Wood, William I.
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: P1176R2
CURRENT APPLICATION NUMBER: US/10/112,267
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
NUMBER OF SEQ ID NOS: 156
SEQ ID NO 110
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1-24
OTHER INFORMATION: Sequence is synthesized.
US-10-112-267-110

Query Match 9.7%; Score 17; DB 15; Length 24;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGCAGGCCAGGCC 93
|||||
DB 17 TCTGCAGGCCAGGCC 1

RESULT 11

US-09-783-590-7073
Sequence 7073, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16.2C1
CURRENT APPLICATION NUMBER: US/09/783,590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7073
LENGTH: 141
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

```
; NAME/KEY: misc feature
; LOCATION: (2)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (67)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (70)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (112)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-7073
```

```
Query Match          9.7%; Score 17; DB 11; Length 141;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      95 TTCTCTTCAGCATGAGA 111
        |||
Db       90 TTCTCTTCAGCATGAGA 106
```

RESULT 12

```
US-09-918-995-12392/c
; Sequence 12392, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12392
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(451)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-12392
```

```
Query Match          9.7%; Score 17; DB 12; Length 451;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 TCCAGGAAGCTTGAGCTT 19
        |||
Db       174 TCCAGGAAGCTTGAGCTT 158
```

RESULT 13

```
US-10-027-632-204519/c
; Sequence 204519, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
```

```
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204519
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-204519
```

```
Query Match          9.7%; Score 17; DB 15; Length 627;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      106 ATGAGAAAGACAAGGGA 122
        |||
Db       457 ATGAGAAAGACAAGGGA 441
```

RESULT 14

```
US-10-027-632-204520/c
; Sequence 204520, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204520
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-204520
```

```
Query Match          9.7%; Score 17; DB 15; Length 627;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      106 ATGAGAAAGACAAGGGA 122
        |||
Db       457 ATGAGAAAGACAAGGGA 441
```

RESULT 15

```
US-10-027-632-204521/c
; Sequence 204521, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
```



```
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204521
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-204521
```

```
Query Match          9.7%; Score 17; DB 15; Length 627;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      106 ATGAGAAAGACAAGGA 122
        ||||||||||||||||
Db      457 ATGAGAAAGACAAGGA 441
```

Search completed: July 28, 2003, 15:36:37
Job time : 48.097 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:16:26 ; Search time 310.442 Seconds
(without alignments)
9129.604 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708
Perfect score: 175
Sequence: 1 AGTCCAGGAACTTGAGCTTT.....GCCTAGAATAACACCCCAA 175

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 segs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	87.4	762	14 BQ195526	BQ195526 UI-R-CN1-
2	35	20.0	315	10 BB374499	BB374499 BB374499
3	35	20.0	369	10 BB792544	BB792544 BB792544
4	35	20.0	418	9 AI225477	AI225477 ue88b01.y
5	35	20.0	792	12 BF138093	BF138093 601785003
6	22	12.6	307	10 BB220676	BB220676 BB220676

C	7	20	11.4	299	12	BF903057	BF903057 CM3-MT019
	8	20	11.4	313	10	BB367824	BB367824 BB367824
	9	20	11.4	538	17	AZ372255	AZ372255 IM0124103
	10	20	11.4	683	17	AG078128	AG078128 Pan trogl
	11	19	10.9	546	10	AW660658	AW660658 99634 MAR
	12	19	10.9	591	17	BH112293	BH112293 RPCI-24-2
	13	19	10.9	932	12	BE778127	BE778127 601463356
	14	19	10.9	957	13	BI334722	BI334722 602999615
	15	19	10.9	1146	17	CNS06NJV	AL406865 T3 end of
	16	19	10.9	5332	17	AF101910	AF101910 AF101910
	17	18	10.3	108	12	BE809197	BE809197 214594 MA
	18	18	10.3	111	10	AW352985	AW352985 35259 MAR
	19	18	10.3	281	10	BB309266	BB309266 BB309266
	20	18	10.3	297	14	F14567	F14567 SSC2H04 Por
	21	18	10.3	320	12	BF561635	BF561635 UI-R-C0-1
	22	18	10.3	367	9	AA804915	AA804915 oa89b02.s
	23	18	10.3	371	10	AV667415	AV667415 AV667415
	24	18	10.3	375	17	AQ134190	AQ134190 HS_3057_A
	25	18	10.3	383	13	BI327290	BI327290 AR076E021
	26	18	10.3	385	12	BF193668	BF193668 245334 MA
	27	18	10.3	386	10	AW654343	AW654343 103751 MA
	28	18	10.3	389	17	AZ277972	AZ277972 RPCI-23-1
	29	18	10.3	395	10	AW977810	AW977810 EST389814
	30	18	10.3	426	9	AA767711	AA767711 ob49b08.s
	31	18	10.3	457	10	AW776806	AW776806 EST335871
	32	18	10.3	461	17	AZ838436	AZ838436 2M0134E10
	33	18	10.3	482	12	BF010204	BF010204 SS95h08.Y
	34	18	10.3	487	12	BF806741	BF806741 IL5-CI014
	35	18	10.3	487	14	BM703250	BM703250 UI-E-CL1-
	36	18	10.3	491	9	AU083781	AU083781 AU083781
	37	18	10.3	491	17	AZ965530	AZ965530 2M0235J07
	38	18	10.3	508	17	AQ430430	AQ430430 HS_5095_A
	39	18	10.3	519	10	BE203038	BE203038 EST403060
	40	18	10.3	530	13	BJ538294	BJ538294 BJ538294
	41	18	10.3	534	13	BJ467272	BJ467272 BJ467272
	42	18	10.3	543	10	AW142440	AW142440 EST292692
	43	18	10.3	560	14	BQ201661	BQ201661 UI-R-DO1-
	44	18	10.3	564	10	BE342161	BE342161 EST395003
	45	18	10.3	627	17	AZ838449	AZ838449 2M0134G11

ALIGNMENTS

RESULT 1
BQ195526/c 762 bp mRNA linear EST 30-APR-2002
LOCUS UI-R-CN1-cmq-k-07-0-UI.s1 UI-R-CN1 Rattus norvegicus cDNA clone
DEFINITION UI-R-CN1-cmq-k-07-0-UI 3', mRNA sequence.
ACCESSION BQ195526 GI:20371077
VERSION BQ195526.1
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 762)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msquares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized cervix library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) Seq primer: M13 Forward POLYA=Yes.

FEATURES

source

1.762
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CNI-cmq-k-07-0-UI"
/clone_1lb="UI-R-CNI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CNI library is a subtracted library derived from the following pool of seven normalized rat libraries: normalized rat seminal vesicles, normalized rat penis, normalized rat bladder, normalized rat cervix, normalized rat brown adipose, normalized rat fundus, and normalized rat salivary gland. It was constructed according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the CNI library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plamid DNA template preparation) comprising: a) a pool of about 34,000 clones from the Rat Unigene Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CA0 and CA1 corresponding to plates R-CA0-AMV through R-CA0-AXS, R-CA0-AZX through R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS, R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKU-K, R-CA0-BKP through R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA, R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA1-BBA through R-CA1-BDA, R-CA1-BHZ through R-CA1-BJF, R-CA1-BJR, R-CA1-BJT through R-CA1-BKB, R-CA1-BKD, R-CA1-BKF, R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLH through R-CA1-BLN, R-CA1-BLS, R-CA1-BLU-V, R-CA1-BNR, and R-CA1-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0, CT0, CU0, CW0, and CX0 and normalized libraries CS0, CT0, CU0, CW0, and CX0 corresponding to plates R-CS0s-GBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVW through R-CW0-BWP, R-CW0-BXN through R-CW0-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CVO and 4,000 from normalized eye library CV1) corresponding to plates R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BVDP (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BFB, R-BV0-ANK through R-BV0-ANR, R-BV0P-AOI through R-BV0P-AOX, and R-BX0-AQY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer

including CY0, CZ0, DA0, DB0, DC0, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYI, R-CZ0-BZB-C, R-DA0-BYJ through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ through R-DB0-BZA, R-DC0-BZI through R-DC0-BZQ, R-DC0-CAY through R-DC0-CBA, R-DD0-BZR through R-DD0-CA, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CNO, that makes up the tracer. The corresponding plates are R-CNO-BKW through R-CNO-BLD, R-CNO-BLG, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT, R-CNO-BLW-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant addresses in the CNO pool corresponding to the following addresses: bkw-a-09-0-UI, bkw-b-09-0-UI, bkw-b-11-0-UI, bkw-b-10-0-UI, bkw-d-01-0-UI, bkw-d-06-0-UI, bkw-g-08-0-UI, bkw-h-12-0-UI, bky-a-05-0-UI, bky-a-06-0-UI, bky-a-11-0-UI, bky-c-06-0-UI, bky-c-09-0-UI, bky-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-e-95-0-UI, bld-1-08-0-UI, bld-f-02-0-UI, blq-h-04-0-UI, blr-a-05-0-UI, blt-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CNO clone (corresponding to the address bky-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG_LIB=UI-R-CNI
TAG_TISSUE=cervix
TAG_SEQ=GACCA"

BASE COUNT 177 a 176 c 198 g 211 t
ORIGIN

Query Match 87.4%; Score 153; DB 14; Length 762;
Best Local Similarity 100.0%; Pred. No. 5.7e-70;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCAGGAGACTGTGATTTTCAGGAATGCACATCTCTTAAGCACTGCCAAAC 60
DB 202 AGTCAGGAGACTGTGATTTTCAGGAATGCACATCTCTTAAGCACTGCCAAAC 143
QY 61 AGGAAGGCTCCACACTCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGAAAGCAAGG 120
DB 142 AGGAAGGCTCCACACTCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGAAAGCAAGG 83
QY 121 GACAGCAGAGTACTCTCTCTGAGGACTAGTC 153
DB 82 GACAGCAGAGTACTCTCTCTGAGGACTAGTC 50

RESULT 2 315 bp mRNA linear EST 13-JUL-2000
BB374499 musculus C130075G16 3' similar to AF126063 Mus musculus
DEFINITION connective tissue growth factor-like protein precursor (Ctgfl) mRNA

, mRNA sequence.

ACCESSION BB374499
VERSION BB374499.1 GI:9086993

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 315)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Horii,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.

cell_line=CCL-142 RAG), (tissue_type=submandibular gland, cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C, cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3), (strain=C3H, tissue_type=brain, cell_line=CRL-1443 BC3H1)"

BASE COUNT 102 a 90 c 86 g 91 t

ORIGIN

Query Match 20.0%; Score 35; DB 10; Length 369;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA 111
|||||
Db 262 TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA 296

RESULT 4
AI225477 418 bp mRNA linear EST 29-OCT-1998
LOCUS ue88b01.y1 Soares_NMPu Mus musculus cDNA clone IMAGE:1498153 5',
DEFINITION mRNA sequence.
ACCESSION AI225477 GI:3808530
VERSION AI225477.1 GI:3808530
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 418)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheilenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:935757
Seq primer: -40RP from Gibco
High quality sequence stop: 403.

FEATURES
Source 1. 418
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1498153"
/clone_lib="Soares_NMPu"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pregnant mouse uterus, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized; library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 121 a 98 c 100 g 99 t

ORIGIN

Query Match 20.0%; Score 35; DB 9; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA 111
|||||
Db 300 TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA 334

RESULT 5
BF138093 792 bp mRNA linear EST 24-OCT-2000
LOCUS 601785003F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4012801 5',
DEFINITION mRNA sequence.
ACCESSION BF138093
VERSION BF138093.1 GI:10977133
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 792)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM9254 row: f column: 02
High quality sequence stop: 705.

FEATURES
Source 1. 792
Location/Qualifiers
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4012801"
/clone_lib="NCI_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 193 a 204 c 205 g 190 t

ORIGIN

Query Match 20.0%; Score 35; DB 12; Length 792;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA 111
|||||
Db 659 TCTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGA 693

RESULT 6
BB220676 307 bp mRNA linear EST 01-JUL-2000
LOCUS BB220676 RIKEN full-length enriched, adult male aorta and vein Mus
DEFINITION musculus cDNA clone A53061F23 3' similar to AF126063 Mus musculus
connective tissue growth factor-like protein precursor (Ctgfl) mRNA
, mRNA sequence.
ACCESSION BB220676
VERSION BB220676.1 GI:8889288
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 307)

AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamana,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.riken.go.jp) for further details.
FEATURES
source
location/Qualifiers
1..307
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A530061F23"
/clone_lib="RIKEN full-length enriched, adult male aorta and vein"
/sex="male"
/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTTAATTAAATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from lambda FLC I."

BASE COUNT 87 a 71 c 70 g 79 t
ORIGIN

Query Match 12.6%; Score 22; DB 10; Length 307;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGCAGGCCACGGCCTTCT 98
|||||
Db 199 TCTGCAGGCCACGGCCTTCT 220
RESULT 7
BF903057/c 299 bp mRNA linear EST 18-JAN-2001
LOCUS CM3-MT0193-151200-565-e02 MT0193 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF903057
ACCESSION BF903057
VERSION BF903057.1 GI:12294516
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 299)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3<2=CM3-MT0193-151200-565-e02&t3=2000-12-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 299.
FEATURES
source
location/Qualifiers
1..299
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MT0193"
/dev_stage="Adult"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 79 a 46 c 55 g 119 t
ORIGIN

Query Match 11.4%; Score 20; DB 12; Length 299;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 TTTTCAGGAATGCACATCTC 43
|||||
Db 80 TTTTCAGGAATGCACATCTC 61

RESULT 8
BB367824
LOCUS
DEFINITION

BB367824 313 bp mRNA linear EST 12-JUL-2000
BB367824 RIKEN full-length enriched, 16 days embryo head Mus musculus cDNA clone C130039K03 3' similar to AF126063 Mus musculus connective tissue growth factor-like protein precursor (Ctgfl) mRNA

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BB367824	BB367824.1	GI:9079653	house mouse. Mus musculus	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.	1 (bases 1 to 313)	Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Saito, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shitagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.	RIKEN Mouse ESTs (Konno, H., et al.).	Unpublished (2000)	
					Contact: Yoshihide Hayashizaki	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute	The Institute of Physical and Chemical Research (RIKEN)	1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan	Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasakawa, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
					Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)	Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.	Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)	Carninci, P. and Hayashizaki, Y.	High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
					Please visit our web site (http://genome.rtc.riken.go.jp) for further details.	Location/Qualifiers	1. 313		
					/organism="Mus musculus"	/strain="C57BL/6J"	/db_xref="taxon:10090"	/clone_1lb="C130039K03"	/clone_1lb="RIKEN full-length enriched, 16 days embryo head"
					/sex="mixed"	/tissue_type="head"	/dev_stage="16 days embryo"	/lab_host="DH10B"	/note="Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
					GAGAGAGAGATCTCGAGTTAATTAATTAATCCCGCCCGCCCGCC 3']	cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'	GAGAGAGAGATCTCGAGTTAATTAATTAATCCCGCCCGCCCGCC 3']	cDNA was cloned into the XhoI and BamHI sites. Vector: a	

modified pBluescript KS(+) after bulk excision from Lambda
FLC I"

BASE COUNT	86 a	78 c	70 g	79 t
ORIGIN				
Query Match	11.4%; Score 20; DB 10; Length 313;			
Best Local Similarity	100.0%; Pred. No. 9.1;			
Matches	20; Conservative	0; Mismatches	0; Indels	0; Gaps

0;

QY 92 CCTTCTCTCAGCATGAGA 111
|||||

Db 220 CCTTCTCTCAGCATGAGA 239

RESULT 9
A2372255

LOCUS 538 bp DNA linear GSS 02-OCT-2000

DEFINITION 1M0124103F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0124103 F, DNA sequence.

ACCESSION A2372255

VERSION A2372255.1 GI:10485955

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0124 row: I column: 03
Seq primer: CGTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 538.
Location/Qualifiers
1..538
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0124103"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gil4732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 162 a 107 c 92 g 177 t

ORIGIN

Query Match 11.4%; Score 20; DB 17; Length 538;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 TTCTCTTCAGCATGAGAAA 113
|||||
Db 310 TTCTCTTCAGCATGAGAAA 329

RESULT 10
AG078128 683 bp DNA linear GSS 03-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-072N12.R, genomic survey sequence.
DEFINITION
ACCESSION AG078128
VERSION AG078128.1 GI:16629930
KEYWORDS GSS.
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC Library clone:PTB-072N12.R.
ORGANISM
REFERENCE
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 683)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbpes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKSI45
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES
Source Location/Qualifiers
1.683
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-072N12.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 198 a 65 c 187 g 226 t 7 others
ORIGIN

Query Match 11.4%; Score 20; DB 17; Length 683;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 AACACGAGAGCTCCACACC 76
|||||
Db 657 AACACGAGAGCTCCACACC 676

RESULT 11
LOCUS AM660658 546 bp mRNA linear EST 25-APR-2001
DEFINITION 99634 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION AM660658
VERSION AM660658.1 GI:7426485
KEYWORDS EST.
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS 1 (bases 1 to 546)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Perte,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keeler,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACAGTCACGACG
Plate: 90 row: L column: 23
Seq primer: ATTAGGTGACACTATAG.
FEATURES
Source Location/Qualifiers
1.546
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 1BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
BASE COUNT 160 a 135 c 147 g 104 t
ORIGIN

Query Match 10.9%; Score 19; DB 10; Length 546;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 GGACAGCAGAGTACTCTCC 138
|||||
Db 205 GGACAGCAGAGTACTCTCC 223

RESULT 12
BH112293/c 591 bp DNA linear GSS 19-JUL-2001
LOCUS RPCI-24-265L18.TV RPCI-24 Mus musculus genomic clone RPCI-24-265L18
DEFINITION
DNA sequence.
ACCESSION BH112293
VERSION BH112293.1 GI:14948266
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 591)
Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akınret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-24
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: RPCI-24-265L18.TJB

Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tdb/bac-ends/mouse/bac_end_intro.html
Plate: 265 row: L column: 18
Seq primer: T7

Class: BAC ends.

FEATURES

source

Location/Qualifiers
1. 591
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="RPCI-24-265L18"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: PTARBAC1; site_1: BamHI; site_2: BamHI; RPCI-24 Mouse BAC library produced by Pieter de Jong. The library was cloned in the PTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 210 a 114 c 121 g 146 t
ORIGIN

Query Match 10.9%; Score 19; DB 17; Length 591;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 GAAGGCTCCACACCTCTGG 81
|||||
Db 59 GAAGGCTCCACACCTCTGG 41

RESULT 13
BE778127/c 932 bp mRNA linear EST 20-OCT-2000
LOCUS 601463356F1 NIH_MGC_67 Homo sapiens CDNA clone IMAGE:3866748 5',
DEFINITION mRNA sequence.
ACCESSION BE778127
VERSION BE778127
KEYWORDS EST.
SOURCE BE778127.1 GI:10199325

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 932)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9611 row: P column: 13
High quality sequence stop: 635.

FEATURES

source

1. 932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3866748"

/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: PCMV-SPORT6; site_1: NotI; site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

BASE COUNT 184 a 197 c 258 g 293 t
ORIGIN

Query Match 10.9%; Score 19; DB 12; Length 932;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 AACACGAGAGGCTCCACAC 75
|||||
Db 859 AACACGAGAGGCTCCACAC 841

RESULT 14
BI334722/c 957 bp mRNA linear EST 30-JUL-2001
LOCUS 602999615F1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:5141931 5',
DEFINITION mRNA sequence.
ACCESSION BI334722
VERSION BI334722
KEYWORDS EST.
SOURCE BI334722.1 GI:15019379

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 957)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1349 row: M column: 04
High quality sequence start: 145
High quality sequence stop: 179.

FEATURES

source

1. 957
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5141931"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: PCMV-SPORT6; site_1: NotI; site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

BASE COUNT 252 a 262 c 193 g 249 t 1 others
ORIGIN

Query Match 10.9%; Score 19; DB 13; Length 957;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 GAAGGCTCCACACCTCTGG 81.
|||||
Db 322 GAAGGCTCCACACCTCTGG 304

RESULT 15
CNS06NJY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:19:16 ; Search time 1764.67 Seconds

(without alignments)
10472.371 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_635

Perfect score: 635
Sequence: 1 GACGCTTCGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 segs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	218	34.3	137964	2 AC126895	AC126895 Rattus no
2	218	34.3	226303	2 AC095418	AC095418 Rattus no
3	166	26.1	137964	2 AC126895	AC126895 Rattus no
4	114	18.0	1741	10 AF259981	AF259981 Rattus no
5	90	14.2	1734	6 AR210324	AR210324 Sequence
6	90	14.2	1734	6 AR210325	AR210325 Sequence
7	90	14.2	1734	10 AF100778	AF100778 Mus muscu
8	90	14.2	61072	10 AL731698	AL731698 Mouse DNA
9	90	14.2	216757	2 AL669906	AL669906 Mus muscu
10	80	12.6	1739	10 AF126063	AF126063 Mus muscu
11	32	5.0	738	6 AR210337	AR210337 Sequence
12	32	5.0	1266	6 AX076919	AX076919 Sequence
13	32	5.0	1266	6 AX464186	AX464186 Sequence
14	32	5.0	1283	9 AF083500	AF083500 Homo sapi
15	32	5.0	1293	6 AR210322	AR210322 Sequence
16	32	5.0	1293	6 AR210323	AR210323 Sequence
17	32	5.0	1309	9 AF074604	AF074604 Homo sapi
18	32	5.0	1427	9 AF100780	AF100780 Homo sapi
19	32	5.0	1450	9 BC017782	BC017782 Homo sapi
20	32	5.0	107260	9 AL139352	AL139352 Human DNA
21	27	4.3	51	6 AR210371	AR210371 Sequence
22	27	4.3	51	6 AX076923	AX076923 Sequence
23	22	3.5	142142	2 AC091173	AC091173 Homo sapi
24	22	3.5	149483	2 AC110904	AC110904 Mus muscu
25	22	3.5	186676	9 AC067881	AC067881 Homo sapi
26	22	3.5	200050	1 AL646068	AL646068 Ralstonia
27	21	3.3	150944	2 AC112856	AC112856 Rattus no
28	21	3.3	154881	2 AC122102	AC122102 Rattus no
29	21	3.3	181343	10 AL671882	AL671882 Mouse DNA
30	21	3.3	215105	2 AC073717	AC073717 Mus muscu
31	20	3.1	1003	5 CHKTCEAA	M73064 Chicken T-c
32	20	3.1	1200	6 AX122384	AX122384 Sequence
33	20	3.1	1481	8 AB033535	AB033535 Oryza sat
34	20	3.1	7172	10 AB008516	AB008516 Mus muscu
35	20	3.1	36296	9 HSE81G9	Z82180 Human DNA s
36	20	3.1	83021	2 AC095541	AC095541 Rattus no
37	20	3.1	84001	9 AL162739	AL162739 Human DNA
38	20	3.1	117840	9 AL162400	AL162400 Human DNA
39	20	3.1	129090	2 AC083961	AC083961 Homo sapi
40	20	3.1	135817	2 AC102328	AC102328 Mus muscu
41	20	3.1	147990	2 AC026144	AC026144 Homo sapi
42	20	3.1	162486	2 AC044779	AC044779 Homo sapi
43	20	3.1	163258	2 AP003514	AP003514 Oryza sat
44	20	3.1	168050	9 AC090526	AC090526 Homo sapi
45	20	3.1	168566	2 AC106167	AC106167 Rattus no

ALIGNMENTS

RESULT 1
AC126895 137964 bp DNA linear HTG 24-JUL-2002
LOCUS
DEFINITION Rattus norvegicus clone CH230-301E4, *** SEQUENCING IN PROGRESS
AC126895
ACCESSION AC126895
VERSION AC126895.1 GI:21724040
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 137964)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,I., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Correll,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssohn,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wöden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 137964)
Worley,K.C.

Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 137964)
Worley,K.C.

Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G2HG
Center clone name: CH230-301E4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89474 bases at least Q40
Consensus quality: 93422 bases at least Q30
Consensus quality: 96506 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1057 1056: contig of 1056 bp in length
1157 1156: gap of unknown length
2335 2335: contig of 1179 bp in length
2336 2435: gap of unknown length
2436 3442: contig of 1007 bp in length
3443 3542: gap of unknown length
3543 5081: contig of 1539 bp in length
5082 5181: gap of unknown length
5182 6424: contig of 1243 bp in length
6425 6524: gap of unknown length
6525 7814: contig of 1290 bp in length
7815 7914: gap of unknown length
7915 9153: contig of 1239 bp in length
9154 9253: gap of unknown length
9154 10445: contig of 1192 bp in length
9254 10445: gap of unknown length
9254 10545: gap of unknown length
10446 11771: contig of 1226 bp in length
10546 11871: gap of unknown length
11772 13240: contig of 1369 bp in length
11872 13340: gap of unknown length
13241 14359: contig of 1019 bp in length
13341 14459: gap of unknown length
14360 15651: contig of 1192 bp in length
14460 15751: gap of unknown length
15652 17494: contig of 1743 bp in length
15752 17594: gap of unknown length
17495 18679: contig of 1085 bp in length
17595 18779: gap of unknown length
18680 20681: contig of 1902 bp in length
18780 20781: gap of unknown length
20682 22118: contig of 1337 bp in length
20782 22218: gap of unknown length
22119 23578: contig of 1360 bp in length
22219 23678: gap of unknown length
23579 25423: contig of 1745 bp in length
23679 25523: gap of unknown length
25424 27808: contig of 2285 bp in length
25524 27908: gap of unknown length
27809 30272: contig of 2364 bp in length
27909 30372: gap of unknown length
30273 33091: contig of 2719 bp in length
30373 33191: gap of unknown length
33092 33191: gap of unknown length
33192 35777: contig of 2586 bp in length
33192 35877: gap of unknown length
35778 36908: contig of 1031 bp in length
35878 37008: gap of unknown length
36909 38930: contig of 1922 bp in length
37009 39030: gap of unknown length
38931 41210: contig of 2180 bp in length
39031 41310: gap of unknown length
41211 42937: contig of 1627 bp in length
41311 43037: gap of unknown length
42938 46120: contig of 3083 bp in length
43038 46220: gap of unknown length
46121 48217: contig of 1997 bp in length
46221 48317: gap of unknown length
48218 50433: contig of 2116 bp in length
48318 50533: gap of unknown length
50434 52880: contig of 2347 bp in length
50534 52980: gap of unknown length
52881 55194: contig of 2214 bp in length
52981 55294: gap of unknown length
55195 57482: contig of 2188 bp in length
55294 57582: gap of unknown length
57482 60781: contig of 3199 bp in length
57583 60881: gap of unknown length
60782 62599: contig of 1718 bp in length
60882

```

* 62600 62699: gap of unknown length
* 62700 65203: contig of 2504 bp in length
* 65204 65303: gap of unknown length
* 65304 66914: contig of 1611 bp in length
* 66915 67014: gap of unknown length
* 67015 70839: contig of 3825 bp in length
* 70840 70939: gap of unknown length
* 70940 75139: contig of 4200 bp in length
* 75140 75239: gap of unknown length
* 75240 78296: contig of 3057 bp in length
* 78297 78396: gap of unknown length
* 78397 83138: contig of 4742 bp in length
* 83139 83238: gap of unknown length
* 83239 88204: contig of 4966 bp in length
* 88205 88304: gap of unknown length
* 88305 92238: contig of 3934 bp in length
* 92239 92338: gap of unknown length
* 92340 97339: contig of 5001 bp in length
* 97340 97439: gap of unknown length
* 97440 103534: contig of 6095 bp in length
* 103535 103634: gap of unknown length
* 103635 107080: contig of 3446 bp in length
* 107081 107180: gap of unknown length
* 107181 115201: contig of 8021 bp in length
* 115202 115301: gap of unknown length
* 115302 121237: contig of 5936 bp in length
* 121238 121337: gap of unknown length
* 121338 127910: contig of 6573 bp in length
* 127911 128010: gap of unknown length
* 128011 137964: contig of 9954 bp in length.

```

```

FEATURES
    source
        1. .137964
            /organism="Rattus norvegicus"
            /db_xref="taxon:10116"
            /clone="CH230-301E4"

```

```

BASE COUNT  31831 a 34027 c 35229 g 31979 t. 4898 others
ORIGIN

```

```

Query Match      34.3%; Score 218; DB 2; Length 137964;
Best Local Similarity 100.0%; Pred. No. 3.2e-106;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY  308 GGTGTGTGCCACACTGTGTCGGACACACCTGTACCTGTCTTGACACACCCAGTGCCC 367
DB  104664 GGTGTGTGCCACACTGTGTCGGACACACCTGTACCTGTCTTGACACACCCAGTGCCC 104723

QY  368 ACAGGGGGTACCCTGTGCTGTGATGGCTGTGCTGTGTAAGTGTGTGCACGAGGCT 427
DB  104724 ACAGGGGGTACCCTGTGCTGTGATGGCTGTGCTGTGTAAGTGTGTGCACGAGGCT 104783

QY  428 GGGGAGTCTCTGCACACCACTGCATGTCTGCGACACCCAGCAGGGCCCTTGTTCAGCC 487
DB  104784 GGGGAGTCTCTGCACACCACTGCATGTCTGCGACACCCAGCAGGGCCCTTGTTCAGCC 104843

QY  488 TGGGGCAGGGCCCTGGCGGCCATGGGGCTGTGTCTCT 525
DB  104844 TGGGGCAGGGCCCTGGCGGCCATGGGGCTGTGTCTCT 104881

```

```

RESULT 2
LOCUS      AC095418                226303 bp      DNA      linear      HTG 11-JUL-2002
DEFINITION Rattus norvegicus clone CH230-7C10, *** SEQUENCING IN PROGRESS ***
ACCESSION  AC095418
VERSION    AC095418.3  GI:21717893
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 226303)

```

AUTHORS

Muzny,D.M., Adams,C., Adio-Obuola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Brannin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Joudah,S., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., Mcleod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Plimms,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N., Slisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vlasov,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

TITLE

JOURNAL
REFERENCE
2 (bases 1 to 226303)

AUTHORS

TITLE

JOURNAL

REFERENCE

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

AUTHORS

TITLE

JOURNAL

REFERENCE

TITLE

JOURNAL

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
* consists of 51 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
1 1051: contig of 1051 bp in length
* 1052 1151: gap of unknown length
* 1152 2286: contig of 1135 bp in length
* 2287 2386: gap of unknown length
* 2387 3433: contig of 1047 bp in length
* 3434 3533: gap of unknown length
* 3534 4614: contig of 1081 bp in length
* 4615 4714: gap of unknown length
* 4715 6581: contig of 1867 bp in length
* 6582 6681: gap of unknown length
* 6682 7773: contig of 1092 bp in length
* 7774 7873: gap of unknown length
* 7874 9530: contig of 1657 bp in length
* 9531 9630: gap of unknown length
* 9631 10661: contig of 1031 bp in length
* 10662 10761: gap of unknown length
* 10762 11958: contig of 1197 bp in length
* 11959 12058: gap of unknown length
* 12059 13140: contig of 1082 bp in length
* 13141 13240: gap of unknown length
* 13241 14777: contig of 1537 bp in length
* 14778 14877: gap of unknown length
* 14878 16063: contig of 1186 bp in length
* 16064 16163: gap of unknown length
* 16164 17585: contig of 1422 bp in length
* 17586 17685: gap of unknown length
* 17686 18916: contig of 1231 bp in length
* 18917 19016: gap of unknown length
* 19017 20305: contig of 1289 bp in length
* 20306 20405: gap of unknown length
* 20406 21537: contig of 1132 bp in length
* 21538 21637: gap of unknown length
* 21638 23703: contig of 2066 bp in length
* 23704 23803: gap of unknown length
* 23804 26306: contig of 2503 bp in length
* 26307 26406: gap of unknown length
* 26407 28431: contig of 2025 bp in length
* 28432 28531: gap of unknown length
* 28532 30615: contig of 2084 bp in length
* 30616 30715: gap of unknown length
* 30716 33215: contig of 2500 bp in length
* 33216 33315: gap of unknown length
* 33316 35357: contig of 2042 bp in length
* 35358 35457: gap of unknown length
* 35458 38012: contig of 2555 bp in length
* 38013 38112: gap of unknown length
* 38113 40338: contig of 2226 bp in length
* 40339 40438: gap of unknown length
* 40439 44051: contig of 3613 bp in length
* 44052 44151: gap of unknown length
* 44152 46294: contig of 2143 bp in length
* 46295 46394: gap of unknown length
* 46395 48845: contig of 2451 bp in length
* 48846 48945: gap of unknown length
* 48946 51753: contig of 2808 bp in length
* 51754 51853: gap of unknown length
* 51854 55915: contig of 4062 bp in length
* 55916 56015: gap of unknown length
* 56016 60468: contig of 4453 bp in length
* 60469 60568: gap of unknown length
* 60569 63640: contig of 3072 bp in length
* 63641 63740: gap of unknown length
* 63741 67888: contig of 4148 bp in length
* 67889 67988: gap of unknown length
* 73014: contig of 5026 bp in length
```

```
* 73015 73114: gap of unknown length
* 73115 77313: contig of 4199 bp in length
* 77314 77413: gap of unknown length
* 77414 83786: contig of 6373 bp in length
* 83787 83886: gap of unknown length
* 83887 89590: contig of 5704 bp in length
* 89591 89690: gap of unknown length
* 89691 93846: contig of 4156 bp in length
* 93847 93946: gap of unknown length
* 93947 98126: contig of 4180 bp in length
* 98127 98226: gap of unknown length
* 98227 104258: contig of 6032 bp in length
* 104259 104358: gap of unknown length
* 104359 110994: contig of 6636 bp in length
* 110995 111094: gap of unknown length
* 111095 119838: contig of 8744 bp in length
* 119839 119938: gap of unknown length
* 119939 126339: contig of 6401 bp in length
* 126340 126439: gap of unknown length
* 126440 134888: contig of 8449 bp in length
* 134889 134988: gap of unknown length
* 134989 142709: contig of 7721 bp in length
* 142710 142809: gap of unknown length
* 142810 149818: contig of 7009 bp in length
* 149819 149918: gap of unknown length
* 149919 158719: contig of 8801 bp in length
* 158720 158819: gap of unknown length
* 158820 169868: contig of 11049 bp in length
* 169869 169968: gap of unknown length
* 169969 180390: contig of 10422 bp in length
* 180391 180490: gap of unknown length
* 180491 191272: contig of 10782 bp in length
* 191273 191372: gap of unknown length
* 191373 209667: contig of 18295 bp in length
* 209668 209767: gap of unknown length
* 209768 226303: contig of 16536 bp in length.
```

FEATURES
source 1. . 226303
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-7C10"

BASE COUNT 57875 a 53353 c 52917 g 57041 t 5117 others

Query Match 34.3%; Score 218; DB 2; Length 226303;
Best Local Similarity 100.0%; Pred. No. 2.9e-106;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 308 GGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGACACCAACCCAGTGCCC 367
    |||||||
Db 147645 GGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGACACCAACCCAGTGCCC 147704

QY 368 ACAGGGGTTACCCCTGTGTCTGTGATGGCTGTGGCTGTGTAAGTGTGTGCACGAGGCT 427
    |||||||
Db 147705 ACAGGGGTTACCCCTGTGTCTGTGATGGCTGTGGCTGTGTAAGTGTGTGCACGAGGCT 147764

QY 428 GGGGAGTCTGTGCGACACCACTGTGATGTGCGACCCAGCCAGGCGCTTGTGCAGCC 487
    |||||||
Db 147765 GGGGAGTCTGTGCGACCACTGTGATGTGCGACCCAGCCAGGCGCTTGTGCAGCC 147824

QY 488 TGGGGCAGGCGCTGTGGGCGCATGGGCTGTGTCTCT 525
    |||||||
Db 147825 TGGGGCAGGCGCTGTGGGCGCATGGGCTGTGTCTCT 147862
```

RESULT 3
AC126895/c 137964 bp DNA linear HTG 24-JUL-2002
LOCUS Rattus norvegicus clone CH230-301E4, *** SEQUENCING IN PROGRESS
DEFINITION *** 49 unordered pieces.
ACCESSION AC126895
VERSION AC126895.1 GI:21724040
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 137964)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsi,F., Howard,S., Huber,J., Huilyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lileu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 137964)
Worley,K.C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 137964)
Worley,K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZHG
Center clone name: CH230-301E4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89474 bases at least Q40
Consensus quality: 93422 bases at least Q30

Consensus quality: 96506 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1056: contig of 1056 bp in length
1057 1156: gap of unknown length
1157 2335: contig of 1179 bp in length
2336 2435: gap of unknown length
2436 3442: contig of 1007 bp in length
3443 3542: gap of unknown length
3543 5081: contig of 1539 bp in length
5082 5181: gap of unknown length
5182 6424: contig of 1243 bp in length
6425 6524: gap of unknown length
6525 7814: contig of 1290 bp in length
7815 7914: gap of unknown length
7915 9153: contig of 1239 bp in length
9154 9253: gap of unknown length
9254 10445: contig of 1192 bp in length
10446 10545: gap of unknown length
10546 11771: contig of 1226 bp in length
11772 11871: gap of unknown length
11872 13240: contig of 1369 bp in length
13241 13340: gap of unknown length
13341 14359: contig of 1019 bp in length
14360 14459: gap of unknown length
14460 15651: contig of 1192 bp in length
15652 15751: gap of unknown length
15752 17494: contig of 1743 bp in length
17495 17594: gap of unknown length
17595 18679: contig of 1085 bp in length
18680 18779: gap of unknown length
18780 20681: contig of 1902 bp in length
20682 20781: gap of unknown length
20782 22118: contig of 1337 bp in length
22119 22218: gap of unknown length
22219 23578: contig of 1360 bp in length
23579 23678: gap of unknown length
23679 25423: contig of 1745 bp in length
25424 25523: gap of unknown length
25524 27808: contig of 2285 bp in length
27809 27908: gap of unknown length
27909 30272: contig of 2364 bp in length
30273 30372: gap of unknown length
30373 33091: contig of 2719 bp in length
33092 33191: gap of unknown length
33192 35777: contig of 2586 bp in length
35778 35877: gap of unknown length
35878 36908: contig of 1031 bp in length
36909 37008: gap of unknown length
37009 38930: contig of 1922 bp in length
38931 39030: gap of unknown length
39031 41210: contig of 2180 bp in length
41211 41310: gap of unknown length
41311 42937: contig of 1627 bp in length
42938 43037: gap of unknown length
43038 46120: contig of 3083 bp in length
46121 46220: gap of unknown length
46221 48217: contig of 1997 bp in length
48218 48317: gap of unknown length
48318 50433: contig of 2116 bp in length
50434 50533: gap of unknown length
50534 52880: contig of 2347 bp in length
52881 52980: gap of unknown length
52981 55194: contig of 2214 bp in length

* 55195 55294: gap of unknown length
* 55295 57482: contig of 2188 bp in length
* 57483 57582: gap of unknown length
* 57583 60781: contig of 3199 bp in length
* 60782 60881: gap of unknown length
* 60882 62599: contig of 1718 bp in length
* 62600 62699: gap of unknown length
* 62700 65203: contig of 2504 bp in length
* 65204 65303: gap of unknown length
* 65304 66914: contig of 1611 bp in length
* 66915 67015 67014: gap of unknown length
* 67015 70839: contig of 3825 bp in length
* 70840 70939: gap of unknown length
* 70940 75139: contig of 4200 bp in length
* 75140 75239: gap of unknown length
* 75240 78296: contig of 3057 bp in length
* 78297 78396: gap of unknown length
* 78397 83138: contig of 4742 bp in length
* 83139 83238: gap of unknown length
* 83239 88204: contig of 4966 bp in length
* 88205 88304: gap of unknown length
* 88305 92238: contig of 3934 bp in length
* 92239 92338: gap of unknown length
* 92339 97439: contig of 5001 bp in length
* 97340 97439: gap of unknown length
* 97440 103534: contig of 6095 bp in length
* 103535 103634: gap of unknown length
* 103635 107080: contig of 3446 bp in length
* 107081 107180: gap of unknown length
* 107181 115201: contig of 8021 bp in length
* 115202 115301: gap of unknown length
* 115302 121237: contig of 5936 bp in length
* 121238 121337: gap of unknown length
* 121338 127910: contig of 6573 bp in length
* 127911 128010: gap of unknown length
* 128011 137964: contig of 9954 bp in length.

FEATURES
source 1.137964
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-301E4"

BASE COUNT 31831 a 34027 c 35229 g 31979 t 4898 others
ORIGIN

Query Match 26.1%; Score 166; DB 2; Length 137964;
Best Local Similarity 99.5%; Pred. No. 3.1e-78;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACGCTTGATCTCCAGAGACCCTGGGTGGAGAGGGCCTTGCAAGGCTGCAGCC 60
Db 74321 GACGCTTGATCTCCAGAGACCCTGGGTGGAGAGGGCCTTGCAAGGCTGCAGCC 74262
QY 61 GCTGGGAGTGGCTTGAATGAGGCTTTATTACTGGGAAGTGAAGAGGCTC 120
Db 74261 GCTGGGAGTGGCTTGAATGAGGCTTTATTACTGGGAAGTGAAGAGGCTC 74202
QY 121 CTGTACAGCTTGCTTAAGCTTAGCAGCTTGCTGGCTTGCGCTTCACACAGTGCAGA 180
Db 74201 CTGTACAGCTTGCTTAAGCTTAGCAGCTTGCTGGCTTGCGCTTCACACAGTGCAGA 74142
QY 181 CACCTTCGTGGTGGCTCCAGGGCCTCACCTTCAGGT 217
Db 74141 CACCTTCGTGGTGGCTCCAGGGCCTCACCTTCAGGT 74105

RESULT 4
LOCUS AF259981 1741 bp mRNA linear ROD 09-MAY-2000
DEFINITION Rattus norvegicus CCN family protein COP-1 (Cop-1) mRNA, complete
ACCESSION AF259981
VERSION AF259981.1 GI:7739780
KEYWORDS

SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1741)
AUTHORS Zhang,R., Averboukh,L., Zhu,W., Zhang,H., Jo,H., Dempsey,P.J.,
Coffey,R.J., Pardee,A.B. and Liang,P.
TITLE Identification of rCop-1, a new member of the CCN protein family,
as a negative regulator for cell transformation
JOURNAL Mol. Cell. Biol. 18 (10), 6131-6141 (1998)
MEDLINE 98414629
PUBMED 9742130
REFERENCE 2 (bases 1 to 1741)
AUTHORS Liang,P.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer
Center, 649 MRB II, Nashville, TN 37232, USA
FEATURES
source 1.1741
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1.1741
/gene="Cop-1"
262.1014
/gene="Cop-1"
/note="secreted protein"
/codon_start=1
/product="CCN family protein COP-1"
/protein_id="AA69011.1"
/db_xref="GI:7739781"
/translation="MRGSPILRLATSLFLCLSMVCAQLCRTPTCTPPTPQCPQGV
LVLDGCGCCKVCARLTLSECEHLHVCEPSGLVCPGAGPGHGAVALDEDDGCEV
NGRRYLDGETFRKNCRLVLCRCDDGGFTCLPLCSEDTVLPWDCPRKRIQVPGKCP
WVCDGVTPIAIOISAQGHOLSLVTPASADAPWPNWSTAWGPCSTTCGLGIATRVSN
QNRFCQLEIQRLCLPRPCLARSHSSWNSAF"

BASE COUNT 386 a 491 c 480 g 384 t
ORIGIN

Query Match 18.0%; Score 114; DB 10; Length 1741;
Best Local Similarity 99.1%; Pred. No. 7e-50;
Matches 214; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 213 CAGGTTGAAGCTGGCTCCACAAGGACACGCTGACATGAGGGCAGCCCACTGATCCAT 272
Db 226 CAGGTTGAAGCTGGCTCCACAAGGACACGCTGACATGAGGGCAGCCCACTGATCCGT 285
QY 273 CTCTGGCCACTTCCTTCTCTGCTTCTCTCAATGCTGTGTGCCAGCTGTGCCGACA 332
Db 286 CTCTGGCCACTTCCTTCTCTGCTTCTCTCAATGCTGTGTGCCAGCTGTGCCGACA 345
QY 333 CCCTGTACCTGTCTTGGACACCAACCCAGTGCCTCCACAGGGGGTACCCTGTGCTGGAT 392
Db 346 CCCTGCACCTGTCTTGGACACCAACCCAGTGCCTCCACAGGGGGTACCCTGTGCTGGAT 405
QY 393 GGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGGCTG 428
Db 406 GGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGGCTG 441

RESULT 5
LOCUS AR210324 1734 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 17 from patent US 6387657.
ACCESSION AR210324
VERSION AR210324.1 GI:21512525
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.

TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 17 14-MAY-2002;
FEATURES Location/Qualifiers
source 1.1734
BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN
Query Match 14.2%; Score 90; DB 6; Length 1734;
Best Local Similarity 100.0%; Pred. No. 5.8e-37;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 410 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCACTGCATGTCTGCGACCCAGCCA 469
Db 418 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCACTGCATGTCTGCGACCCAGCCA 477
QY 470 GGGCCTGTTGTGTCAGCCTGGGGCAGGCC 499
Db 478 GGGCCTGTTGTGTCAGCCTGGGGCAGGCC 507
RESULT 6
AR210325/c 1734 bp DNA linear PAT 20-JUN-2002
LOCUS AR210325 Sequence 18 from patent US 6387657.
DEFINITION AR210325
ACCESSION AR210325
VERSION AR210325.1 GI:21512526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 18 14-MAY-2002;
FEATURES Location/Qualifiers
source 1.1734
BASE COUNT 393 a 495 c 491 g 355 t
ORIGIN
Query Match 14.2%; Score 90; DB 6; Length 1734;
Best Local Similarity 100.0%; Pred. No. 5.8e-37;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 410 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCACTGCATGTCTGCGACCCAGCCA 469
Db 1317 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCACTGCATGTCTGCGACCCAGCCA 1258
QY 470 GGGCCTGTTGTGTCAGCCTGGGGCAGGCC 499
Db 1257 GGGCCTGTTGTGTCAGCCTGGGGCAGGCC 1228
RESULT 7
AF100778 1734 bp mRNA linear ROD 17-DEC-1998
LOCUS AF100778
DEFINITION Mus musculus connective tissue growth factor related protein WISP-2
(Wisp2) mRNA, complete cds.
ACCESSION AF100778
VERSION AF100778.1 GI:4028578
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1734)
AUTHORS Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A.,
Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C.,
Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D.,
Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.
TITLE WISP genes are members of the connective tissue growth factor

family that are up-regulated in wnt-1-transformed cells and
aberrantly expressed in human colon tumors
Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)
MEDLINE 99061933
PUBMED 9843955
REFERENCE 2 (bases 1 to 1734)
AUTHORS Pennica,D.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA
Way, South San Francisco, CA 94080, USA
FEATURES Location/Qualifiers
source 1.1734
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="C57MG"
/cell_type="epithelial"
/tissue_type="mammary"
/note="transformed by wnt-1"
1.1734
/gene="Wisp2"
257.1012
/gene="Wisp2"
/codon_start=1
/product="connective tissue growth factor related protein
WISP-2"
/protein_id="AAC96320.1"
/db_xref="GI:4028579"
/translation="MRGNPLIHLAISFLCILSMVYSQLCPAPACPMTPPCPGVP
LVLDGCGCCRVCAARLGESECDLHVCDPSQGLVCQPGAGPSGRGAVCLFEEDGSCV
NGRRVLDGETFEFKPNCRLVLCRCDDGFTCLPLCSEDEVRLPSWDCPRPRRIOVGRCP
WVCDQAVMQPAIQPSAQQHQLSALVTPASADGPCPNWSTAMGPCSTTCGLGIATRV
NONRFOLEIQRLCLSRPCLASRSWSNSAF"
BASE COUNT 355 a 491 c 495 g 393 t
ORIGIN
Query Match 14.2%; Score 90; DB 10; Length 1734;
Best Local Similarity 100.0%; Pred. No. 5.8e-37;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 410 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCACTGCATGTCTGCGACCCAGCCA 469
Db 418 AGTGTGTCACGAGGCTGGGGAGTCTCGACACCACTGCATGTCTGCGACCCAGCCA 477
QY 470 GGGCCTGTTGTGTCAGCCTGGGGCAGGCC 499
Db 478 GGGCCTGTTGTGTCAGCCTGGGGCAGGCC 507
RESULT 8
AL731698 61072 bp DNA linear ROD 24-MAY-2002
LOCUS AL731698
DEFINITION Mouse DNA sequence from clone RP23-161B3 on chromosome 2, complete
sequence.
ACCESSION AL731698
VERSION AL731698.10 GI:21214309
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 61072)
AUTHORS Wallis,J.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 25, 2002 this sequence version replaced gi:21213601.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-161B3 is from the RPCI-23 mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES
source
1. 61072
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-161B3"
/clone_lib="RPCI-23"
BASE COUNT 15405 a 15368 c 15033 g 15266 t
ORIGIN
Query Match 14.2%; Score 90; DB 10; length 61072;
Best Local Similarity 100.0%; Pred. No. 2.9e-37;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 410 AGTGTGTGCACGAGGCTGGGGAGTCTGCGACCACTGCATGTCTGCAGCCCGACCA 469
|||||
Db 10029 AGTGTGTGCACGAGGCTGGGGAGTCTGCGACCACTGCATGTCTGCAGCCCGACCA 10088
QY 470 GGGCCTGCTTGTACGCTGGGGCAGGCC 499
|||||
Db 10089 GGGCCTGCTTGTACGCTGGGGCAGGCC 10118
RESULT 9
AL669906 216757 bp DNA linear HTG 24-JUL-2002
LOCUS
DEFINITION Mus musculus chromosome 2 clone RP23-217C2, *** SEQUENCING IN
PROGRESS ***, 21 unordered pieces.
ACCESSION AL669906
VERSION AL669906.5 GI:21955520
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 216757)
JOURNAL Sims, S.
COMMENT Direct Submission
Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 25, 2002 this sequence version replaced gi:18181793.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: bm217C2
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 21865 bases at least Q40
Consensus quality: 213403 bases at least Q30
Consensus quality: 214139 bases at least Q20
Insert size: 214757; sum-of-contigs
Insert size: 234243; 1.8% error; agarose-ff

Quality coverage: 6.29x in Q20 bases; sum-of-contigs. Quality
coverage: 5.90x in Q20 bases; agarose-ff

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4000: contig of 4000 bp in length
* 4001 4100: gap of 100 bp
* 4101 15267: contig of 11167 bp in length
* 15268 15367: gap of 100 bp
* 15368 34970: contig of 19603 bp in length
* 34971 35070: gap of 100 bp
* 35071 39671: contig of 4601 bp in length
* 39672 39771: gap of 100 bp
* 39772 46360: contig of 6589 bp in length
* 46361 46460: gap of 100 bp
* 46461 49809: contig of 3349 bp in length
* 49810 49909: gap of 100 bp
* 49910 54509: contig of 4600 bp in length
* 54510 54609: gap of 100 bp
* 54610 65989: contig of 11380 bp in length
* 65990 66089: gap of 100 bp
* 66090 84635: contig of 18546 bp in length
* 84636 84735: gap of 100 bp
* 84736 87399: contig of 2664 bp in length
* 87400 87499: gap of 100 bp
* 87500 98601: contig of 11102 bp in length
* 98602 98701: gap of 100 bp
* 98702 105522: contig of 6821 bp in length
* 105523 105622: gap of 100 bp
* 105623 119773: contig of 14151 bp in length
* 119774 119873: gap of 100 bp
* 119874 134552: contig of 14679 bp in length
* 134553 134652: gap of 100 bp
* 134653 141202: contig of 6550 bp in length
* 141203 141302: gap of 100 bp
* 141303 145808: contig of 4506 bp in length
* 145809 145908: gap of 100 bp
* 145909 151572: contig of 5664 bp in length
* 151573 151672: gap of 100 bp
* 151673 169686: contig of 18014 bp in length
* 169687 169786: gap of 100 bp
* 169787 200095: contig of 30309 bp in length
* 200096 200195: gap of 100 bp
* 200196 213877: contig of 13682 bp in length
* 213878 213977: gap of 100 bp
* 213978 216757: contig of 2780 bp in length.
FEATURES
source
1. 216757
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-217C2"
/clone_lib="RPCI-23"
1. 4000
/note="assembly_fragment:03429
clone_end:SP6
vector_side:left"
misc_feature
4101. 15267
/note="assembly_fragment:00814
fragment_chain:1"
misc_feature
15368. 34970
/note="assembly_fragment:00812
fragment_chain:1"
misc_feature
35071. 39671
/note="assembly_fragment:03544
fragment_chain:1"
misc_feature
39772. 46360

misc_feature /note="assembly_fragment:01438
fragment_chain:1"
46461..49809
/note="assembly_fragment:03410
fragment_chain:2"
49910..54509
/note="assembly_fragment:00811
fragment_chain:2"
54610..65989
/note="assembly_fragment:02217
fragment_chain:2"
66090..84635
/note="assembly_fragment:03529
fragment_chain:2"
84736..87399
/note="assembly_fragment:01019
fragment_chain:3"
87500..98601
/note="assembly_fragment:01170
fragment_chain:3"
98702..105522
/note="assembly_fragment:02829
fragment_chain:3"
105623..119773
/note="assembly_fragment:01476
fragment_chain:4"
119874..134552
/note="assembly_fragment:00505
fragment_chain:4"
134653..141202
/note="assembly_fragment:01951
fragment_chain:4"
141303..145808
/note="assembly_fragment:02637
fragment_chain:5"
145909..151572
/note="assembly_fragment:02352
fragment_chain:5"
151673..169686
/note="assembly_fragment:02207
fragment_chain:5"
169787..200095
/note="assembly_fragment:02849
fragment_chain:6"
200196..213877
/note="assembly_fragment:01052
fragment_chain:6"
213978..216757
/note="assembly_fragment:01318
clone_end:T7
vector_side:right"
BASE COUNT 56425 a 51573 c 51904 g 54852 t 2003 others
ORIGIN
Query Match 14.2%; Score 90; DB 2; Length 216757;
Best Local Similarity 100.0%; Pred.No. 2.3e-37;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 410 AGTGTGACGAGAGCTGGGGAGTCTGCACACCTGCATGTGCGACCCAGCCA 469
|||||
DB 166450 AGTGTGACGAGAGCTGGGGAGTCTGCACACCTGCATGTGCGACCCAGCCA 166509
QY 470 GGGCTGTGTCTACGCTGGGCGAGCCC 499
|||||
DB 166510 GGGCTGTGTCTACGCTGGGCGAGCCC 166539

RESULT 10
AF126063 1739 bp mRNA linear ROD 12-OCT-1999
LOCUS AF126063
DEFINITION Mus musculus connective tissue growth factor-like protein precursor
(Ctgf1) mRNA, complete cds.
ACCESSION AF126063

VERSION AF126063.1 GI:4337059
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 1739)
AUTHORS Kumar,S., Hand,A.T., Connor,J.R., Dodds,R.A., Ryan,P.J.,
Trill,J.J., Fisher,S.M., Nuttall,M.E., Lipschutz,D.B., Zou,C.,
Hwang,S.M., Volta,B.J., James,I.E., Rleman,D.J., Gowen,M. and
Lee,J.C.
TITLE Identification and cloning of a connective tissue growth
factor-like cDNA from human osteoblasts encoding a novel regulator
of osteoblast functions
JOURNAL J. Biol. Chem. 274 (24), 17123-17131 (1999)
MEDLINE 99287915
PUBMED 10358067
REFERENCE 2 (bases 1 to 1739)
AUTHORS Kumar,S. and Zou,C.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109,
Smithkline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406,
USA

FEATURES
source Location/Qualifiers
1..1739
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="lung"
1..1739
/gene="Ctgf1"
242..997
/gene="Ctgf1"
/note="similar to the Mus musculus WISP-2 protein encoded
by the sequence presented in GenBank Accession Number
AF100778; putative growth factor; CTGF-L; contains IGF
binding (IGFBP), Von Willebrand Factor type C (VWC) repeat
and thrombospondin type I (TSP1) domains; member of the
CCN (CTGF/Cyrl/Nov) family; lacks the fourth
carboxy-terminal (CT) domain present in other members of
the CCN family"
/codon_start=1
/product="connective tissue growth factor-like protein
precursor"
/protein_id="AAD18058.1"
/db_xref="GI:4337060"
/translation="MRGNPLIHLLAISFLCILSMVYSQICPAPCACPWTPPOCPGVP
LVLDGCCRCRVCAARLIGESCDLHVCDPSQGLVCQPGAPSGRGAVCLFEEDGSGEV
NGRRYLIDGETFKPNCRVLCRCDGFTCLPLCSEDRPSWDCPRPRRIQVGRCPPE
WYCDQAVMQPAIQPSAQGHQLSALVTPASADGPCPNWSTAWGPCSTTCGLGIATRVS
NONRRCOLEIQRLCLSRPCLASRSRSHGWSNAF"

BASE COUNT 375 a 480 c 489 g 395 t
ORIGIN
Query Match 12.6%; Score 80; DB 10; Length 1739;
Best Local Similarity 100.0%; Pred.No. 1.4e-31;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 420 CGGAGCTGGGGAGTCTGCACACCTGCATGTCTGCGACCCAGGCCCTGTT 479
|||||
DB 413 CGGAGCTGGGGAGTCTGCACACCTGCATGTCTGCGACCCAGGCCCTGTT 472
QY 480 TGTACGCTGGGGCGAGCCC 499
|||||
DB 473 TGTACGCTGGGGCGAGCCC 492

RESULT 11
AR210337 738 bp DNA linear PAT 20-JUN-2002
LOCUS AR210337
DEFINITION Sequence 38 from patent US 6387657.
ACCESSION AR210337
VERSION AR210337.1 GI:21512542
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 738)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 38 14-MAY-2002;
FEATURES Location/Qualifiers
source 1..738
BASE COUNT 104 a 272 c 238 g 124 t
ORIGIN

Query Match 5.0%; Score 32; DB 6; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
Db 115 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 146

RESULT 12
LOCUS AX076919 1266 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 31 from Patent WO0105836.
ACCESSION AX076919
VERSION AX076919.1 GI:13121575
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1266)
AUTHORS Botstein,D., Goddard,A., Gurney,A.L., Hillan,K.J., Roy,M.A. and
Wood,W.I.
TITLE Polypeptidic compositions and methods for the treatment of tumors
JOURNAL Patent: WO 0105836-A 31 25-JAN-2001;
FEATURES Location/Qualifiers
source 1..1266
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 216 a 418 c 390 g 242 t
ORIGIN

Query Match 5.0%; Score 32; DB 6; Length 1266;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
Db 136 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 167

RESULT 13
LOCUS AX464186 1266 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 319 from Patent WO0140466.
ACCESSION AX464186
VERSION AX464186.1 GI:21899109
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Baker,K.P., Beresini,M., DeForge,L., Desnoyers,L., Filvaroff,E.,
Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
Wood,W.L. and Zhang,Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding

JOURNAL same
Patent: WO 0140466-A 319 07-JUN-2001;
Genentech Inc. (US)
FEATURES Location/Qualifiers
source 1..1266
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 216 a 418 c 390 g 242 t
ORIGIN

Query Match 5.0%; Score 32; DB 6; Length 1266;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
Db 136 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 167

RESULT 14
LOCUS AF083500 1283 bp mRNA linear PRI 04-NOV-1998
DEFINITION Homo sapiens connective tissue growth factor-like protein
precursor, mRNA, complete cds.
ACCESSION AF083500
VERSION AF083500.1 GI:3462835
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1283)
AUTHORS Kumar,S., Hand,A.T., Connor,J.C., Dodds,R.A., Ryan,P.J.,
Trill,J.G., Fisher,S.M., Slemmon,J.R., Lipshutz,D.B.,
Bartholomew,V., James,I.E., Rleman,D.J., Gowen,M. and lee,J.C.
TITLE Identification and cloning of CTGF-L from human osteoblasts, a
novel cysteine rich protein containing an IGF binding domain
JOURNAL Bone 23 (5), S240 (1998)
REFERENCE 2 (bases 1 to 1283)
AUTHORS Kumar,S.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1998) Bone & Cartilage Biology, UW 2109,
Smithkline Beecham, 709 Swedeland Road, King of Prussia, PA 19406,
USA
FEATURES Location/Qualifiers
source 1..1283
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q12-q13"
/cell_type="primary osteoblast"
9..761
/note="CTGF-L; encodes IGF binding (IGFB), Von Willebrand
factor type C (VWC) and thrombospondin type I (TSPI)
domains; member of the CCN (CTGF/cyrl61/nov) family; lacks
the fourth carboxy-terminal domain present in other
members of the CCN family"
/codon_start=1
/product="connective tissue growth factor-like protein
precursor"
/protein_id="AAC70350.1"
/db_xref="GI:3462836"
/translation="MRGTPKTHLLAFLSLCLLSKVRITQCPPTCPWPPRCPPLGVP
LVLDGCGCCRCVRCARLGEPCDQLHVCDAQGLVCPGAGGRCALCLLAEDSSCEV
NGRLYREGETFQPHCSIRCRCEDEGFTCVPLCSQDVRLPSWDCPHPRVEVLGKCPE
WVCGGGGLGTQPLPAQGPQFSGLVSLPPGVPCPEWSTAWGPCSTTCGLGMATRVSN
QNRFCRLDTQRLCLSRPCPSGRSPQNSAF"
BASE COUNT 235 a 418 c 389 g 241 t
ORIGIN

Query Match 5.0%; Score 32; DB 9; Length 1283;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGGTCTGATGGCTGTGCTG 406
Db 135 GTACCCCTGGTCTGATGGCTGTGCTG 166

RESULT 15
AR210322
LOCUS AR210322 1293 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 13 from patent US 6387657.
ACCESSION AR210322
VERSION AR210322.1 GI:21512523
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1293)
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 13 14-MAY-2002;
FEATURES Location/Qualifiers
source 1..1293
/organism="Unknown"

BASE COUNT 232 a 425 c 393 g 243 t
ORIGIN

Query Match 5.0%; Score 32; DB 6; Length 1293;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGGTCTGATGGCTGTGCTG 406
Db 148 GTACCCCTGGTCTGATGGCTGTGCTG 179

Search completed: July 28, 2003, 21:14:43
Job time : 1766.67 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:31:11 ; Search time 166.775 Seconds
(without alignments)
8574.520 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_635
Perfect score: 635
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GTGATGACGGTGCTTCACC 635

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 segs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	635	100.0	1708	AAZ07516	Rat HICP polypepti
2	387	60.9	753	AAZ07517	Rat HICP polypepti
3	318	50.1	681	AAZ07521	Rat HICP mature po
4	210	33.1	210	AAZ07519	Rat HICP IGFBP dom
5	90	14.2	177	AAZ07518	Rat HICP VWC domai
6	90	14.2	753	AAZ07518	Mouse WISP-2 prote
7	90	14.2	1734	AAZ07518	Mouse WISP-2 prote
8	32	5.0	738	AAZ07518	Human WISP-2 prote
9	32	5.0	750	AAZ07518	Human WISP-2 prote

10	32	5.0	1257	20	AAZ07516	EGF-like homologue
11	32	5.0	1266	21	AAZ07516	Human PRO261 nucle
12	32	5.0	1266	22	AAZ07516	Human CDNA sequenc
13	32	5.0	1266	22	AAZ07516	PRO261 coding sequ
14	32	5.0	1266	22	AAZ07516	Human angiogenesis
15	32	5.0	1267	21	AAZ07516	Human PRO261 CDNA
16	32	5.0	1285	19	AAZ07516	Human connective t
17	32	5.0	1293	20	AAZ07516	Human WISP-2 prote
18	32	5.0	1309	22	AAZ07516	Connective tissue
19	32	5.0	1337	22	AAZ07516	Human secreted pro
20	32	5.0	1352	22	AAZ07516	Human secreted pro
21	32	5.0	1522	20	AAZ07516	Human growth facto
22	32	5.0	2136	22	AAZ07516	Human full-length
23	32	5.0	13255	22	AAZ07516	Human immune/haema
24	27	4.3	51	20	AAZ07516	Human WISP-2 probe
25	27	4.3	51	21	AAZ07516	Human PRO261 hybr1
26	27	4.3	51	21	AAZ07516	Probe for PRO261 n
27	27	4.3	51	22	AAZ07516	PRO261 probe #1.
28	27	4.3	51	22	AAZ07516	Human PRO261 hybr1
29	23	3.6	634	22	AAZ07516	Human CDNA 5'-end
30	23	3.6	634	22	AAZ07516	Human CDNA clone r
31	20	3.1	1200	22	AAZ07516	C glutamicum codin
32	20	3.1	34980	22	AAZ07516	C glutamicum codin
33	19	3.0	203	24	AAZ07516	Human ORFX polynuc
34	19	3.0	217	24	AAZ07516	Corn tassal-derive
35	19	3.0	318	22	AAZ07516	Human breast cance
36	19	3.0	332	22	AAZ07516	Human breast cance
37	19	3.0	372	18	AAZ07516	Human host cell pr
38	19	3.0	372	22	AAZ07516	Human host cell pr
39	19	3.0	405	22	AAZ07516	Human protein enco
40	19	3.0	425	17	AAZ07516	TNF-R p55IC/Fas-IC
41	19	3.0	443	22	AAZ07516	Human breast cancer
42	19	3.0	616	21	AAZ07516	Human colon cancer
43	19	3.0	816	24	AAZ07516	Human sumo-1 DNA.
44	19	3.0	816	24	AAZ07516	Human CDNA differe
45	19	3.0	816	24	AAZ07516	Kidney cancer rela

ALIGNMENTS

RESULT 1	
AAZ07516	
ID	AAZ07516 standard; CDNA; 1708 BP.
XX	
AC	AAZ07516;
XX	
DT	26-NOV-1999 (first entry)
XX	
DE	Rat HICP polypeptide encoding CDNA.
XX	
KW	Heparin-induced CGN-like protein; HICP; cell-associated activity; ss;
KW	cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX	
OS	Rattus sp.
XX	
PN	W09947556-A2.
XX	
PD	23-SEP-1999.
XX	
PF	18-MAR-1999; 99WO-US05999.
XX	
PR	19-MAR-1998; 98US-0044273.
XX	
PA	(TUFT) TUFTS COLLEGE.
XX	
PI	Castellot JJ;
XX	
DR	WPI; 1999-562060/47.
DR	P-PSDB; AAY27434.
XX	
PT	Nucleic acid sequences encoding rat heparin-induced CGN-like protein, used in methods to identify modulators or in diagnostic applications
PT	


```
XX Claim 2; Fig 1; 108bp; English.
XX
CC This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein.
CC Agents that stimulate or inhibit HICP protein activity or expression,
CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to
CC modulate cell-associated activity. HICP modulators can be used to treat
CC disorders characterized by aberrant HICP protein activity or expression.
CC Probes capable of hybridizing to HICP mRNA or antibodies specific for
CC HICP can be used to detect HICP activity in a biological sample. HICP
CC can be used to treat disorders, such as a cardiovascular or fibrotic
CC disorder, characterized by aberrant cell proliferation.
XX
SQ Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Query Match      100.0%; Score 635; DB 20; Length 1708;
Best Local Similarity 100.0%; Pred. No. 8.3e-301;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGCTTCTGATCTCCAGAGGACCCTGGGGTGAGACAGGGCCCTGGCAAGCTGCAGCC 60
   |||||||
Db 1 GACGCTTCTGATCTCCAGAGGACCCTGGGGTGAGACAGGGCCCTGGCAAGCTGCAGCC 60

OY 61 GCTGGGAGTGGCTTGAATGAGGCTTTATTAAGTGGAACTGAGAGCTAAGAGCTC 120
   |||||||
Db 61 GCTGGGAGTGGCTTGAATGAGGCTTTATTAAGTGGAACTGAGAGCTAAGAGCTC 120

OY 121 CTGTACAGCTTGTCTTAAAGTCTTAGACACTTGTGGGCTTGGGCTTACACACTGTGAGA 180
   |||||||
Db 121 CTGTACAGCTTGTCTTAAAGTCTTAGACACTTGTGGGCTTGGGCTTACACACTGTGAGA 180

OY 181 CACCTTCTGTGGTGGCTCCACGGGCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGGAC 240
   |||||||
Db 181 CACCTTCTGTGGTGGCTCCACGGGCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGGAC 240

OY 241 ACGGTGACATGAGGGGACGCCCACTGATCCATCTTCTGGGCACCTCCTTCTGCTTC 300
   |||||||
Db 241 ACGGTGACATGAGGGGACGCCCACTGATCCATCTTCTGGGCACCTCCTTCTGCTTC 300

OY 301 TCTCAATGCTGTGCTGCCCAAGCTGTGCCGACACCCCTGTAGCTGTCTTGACACCACCCC 360
   |||||||
Db 301 TCTCAATGCTGTGCTGCCCAAGCTGTGCCGACACCCCTGTAGCTGTCTTGACACCACCCC 360

OY 361 AGTCCCACAGGGGGTACCCCTGTGTGTGATGGCTGTGGCTGTAAAGTGTGTGCAC 420
   |||||||
Db 361 AGTCCCACAGGGGGTACCCCTGTGTGTGATGGCTGTGGCTGTAAAGTGTGTGCAC 420

OY 421 GGAGGCTGGGAGTCTCTCGACCACTGATCTGTGCAGACCCAGCCAGGCTGTGTT 480
   |||||||
Db 421 GGAGGCTGGGAGTCTCTCGACCACTGATCTGTGCAGACCCAGCCAGGCTGTGTT 480

OY 481 GTCAGCCTGGGGAGGCCCTGGGGCCATGGGGCTGTGTCTCTTGATGAGATGACG 540
   |||||||
Db 481 GTCAGCCTGGGGAGGCCCTGGGGCCATGGGGCTGTGTCTCTTGATGAGATGACG 540

OY 541 GTAGCTGTGAGGTGAATGGCCGAGGTACCTGATGAGAGACCTTTAAACCAATTGCA 600
   |||||||
Db 541 GTAGCTGTGAGGTGAATGGCCGAGGTACCTGATGAGAGACCTTTAAACCAATTGCA 600

OY 601 GGGTCTGTGGCCGCTGTGATGACGGGTGGCTTCACC 635
   |||||||
Db 601 GGGTCTGTGGCCGCTGTGATGACGGGTGGCTTCACC 635

RESULT 2
AAZ07517 standard; cDNA; 753 BP.
XX
AC AAZ07517;
XX
DT 26-NOV-1999 (first entry)
XX
DE Rat HICP polypeptide coding sequence.
```

```
XX Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX
OS Rattus sp.
XX
PN W09947556-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT ) TUFTS COLLEGE.
XX
PI Castellet JJ;
XX
DR WPI; 1999-562060/47.
DR P-PSDB; AAY27434.
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications
XX
PS Claim 5; Fig 1; 108bp; English.
XX
CC The invention provides a rat heparin-induced CCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents the coding sequence of rat HICP.
XX
SQ Sequence 753 BP; 132 A; 235 C; 230 G; 156 T; 0 other;

Query Match      60.9%; Score 387; DB 20; Length 753;
Best Local Similarity 100.0%; Pred. No. 2.2e-179;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 ATGAGGGGACGCCCACTGATCCATCTTCTGGCACTTCTCTTCTGCTCTCTCAATG 308
   |||||||
Db 1 ATGAGGGGACGCCCACTGATCCATCTTCTGGCACTTCTCTTCTGCTCTCTCTCAATG 60

OY 309 GTGTGTGCCAGCTGTGCCGGACACCCCTGTACTGTCTTGGACACACCCAGTGCCCA 368
   |||||||
Db 61 GTGTGTGCCAGCTGTGCCGGACACCCCTGTACTGTCTTGGACACACCCAGTGCCCA 120

OY 369 CAGGGGTACCCCTGTGTCTGTGATGGCTGTGGCTGTCTTAAGTGTGTGACGGAGCTG 428
   |||||||
Db 121 CAGGGGTACCCCTGTGTCTGTGATGGCTGTGGCTGTCTTAAGTGTGTGACGGAGCTG 180

OY 429 GGGGAGTCTGCGACCACTGATGTGTGCGACCCCAAGCCAGGGCTGTGTGACGCT 488
   |||||||
Db 181 GGGGAGTCTGCGACCACTGATGTGTGCGACCCCAAGCCAGGGCTGTGTGACGCT 240

OY 489 GGGGAGGGCCCTGGGGCCATGGGGCTGTGTCTCTTGATGAGGATGACGGTAGCTGT 548
   |||||||
Db 241 GGGGAGGGCCCTGGGGCCATGGGGCTGTGTCTCTTGATGAGGATGACGGTAGCTGT 300

OY 549 GAGGTGAATGGCCGAGGTACCTGATGAGAGACCTTTAAACCAATTGACGGTCTG 608
   |||||||
Db 301 GAGGTGAATGGCCGAGGTACCTGATGAGAGACCTTTAAACCAATTGACGGTCTG 360

OY 609 TGCCGCTGTGATGACGGGTGGCTTCACC 635
   |||||||
Db 361 TGCCGCTGTGATGACGGGTGGCTTCACC 387

RESULT 3
```

AAZ07521
ID AAZ07521 standard; cDNA; 681 BP.
XX
AC AAZ07521;
XX
DT 26-NOV-1999 (first entry)
XX
DE Rat HICP mature polypeptide coding sequence.
XX
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.
XX
OS Rattus sp.
XX
PN WO9947556-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Castellot JJ;
XX
DR WPI; 1999-562060/47.
DR P-PSDB; AAY27440.
XX
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications -
XX
XX
PS Disclosure; Fig 2; 108bp; English.
XX
CC The invention provides a rat heparin-induced CCN-like protein (HICP).
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents the coding sequence of rat HICP mature
CC polypeptide.
XX
SQ Sequence 681 BP; 121 A; 211 C; 215 G; 134 T; 0 other;

Query Match 50.1%; Score 318; DB 20; Length 681;
Best Local Similarity 100.0%; Pred. No. 1.3e-145;
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGCCACAGGGGTA 377
DB 1 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGCCACAGGGGTA 60
QY 378 CCCCTGTGCTGGATGGCTGTGGCTGTGAAGTGTGTGCACGGAGGCTGGGGAGTCC 437
DB 61 CCCCTGTGCTGGATGGCTGTGGCTGTGAAGTGTGTGCACGGAGGCTGGGGAGTCC 120
QY 438 TGGGACCACTGCATGTCTGCGACCCGACCCAGGGCCTGTTGTTCAGCCTGGGGCAGGC 497
DB 121 TGGGACCACTGCATGTCTGCGACCCGACCCAGGGCCTGTTGTTCAGCCTGGGGCAGGC 180
QY 498 CCTGGCGCCATGGGGCTGTGTCTCTTTGGATGAGATGACGGTAGCTGTGAGTGAAT 557
DB 181 CCTGGCGCCATGGGGCTGTGTCTCTTTGGATGAGATGACGGTAGCTGTGAGTGAAT 240
QY 558 GGGCGCAGTACCTGGATGAGAGAGACCTTTAAACCAATTCAGAGGTCCTGTGCCGCTGT 617
DB 241 GGGCGCAGTACCTGGATGAGAGAGACCTTTAAACCAATTCAGAGGTCCTGTGCCGCTGT 300
QY 618 GATGACGGTGGCTTCACC 635

DB 301 GATGACGGTGGCTTCACC 318
|||||

RESULT 4
ID AAZ07519 standard; cDNA; 210 BP.
XX
AC AAZ07519;
XX
DT 26-NOV-1999 (first entry)
XX
DE Rat HICP IGFBP domain encoding cDNA.
XX
DE
XX
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder;
KW insulin-like growth factor binding protein; IGFBP.
XX
OS Rattus sp.
XX
PN WO9947556-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Castellot JJ;
XX
DR WPI; 1999-562060/47.
DR P-PSDB; AAY27438.
XX
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications -
XX
XX
PS Disclosure; Page 104; 108bp; English.
XX
CC The invention provides a rat heparin-induced CCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents a nucleotide sequence encoding the insulin-
CC like growth factor binding protein (IGFBP) domain of HICP polypeptide.
XX
SQ Sequence 210 BP; 27 A; 65 C; 74 G; 44 T; 0 other;

Query Match 33.1%; Score 210; DB 20; Length 210;
Best Local Similarity 100.0%; Pred. No. 9.9e-93;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGCCACAGGGGTA 377
DB 1 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGCCACAGGGGTA 60
QY 378 CCCCTGTGCTGGATGGCTGTGGCTGTGAAGTGTGTGCACGGAGGCTGGGGAGTCC 437
DB 61 CCCCTGTGCTGGATGGCTGTGGCTGTGAAGTGTGTGCACGGAGGCTGGGGAGTCC 120
QY 438 TGGGACCACTGCATGTCTGCGACCCGACCCAGGGCCTGTTGTTCAGCCTGGGGCAGGC 497
DB 121 TGGGACCACTGCATGTCTGCGACCCGACCCAGGGCCTGTTGTTCAGCCTGGGGCAGGC 180
QY 498 CCTGGCGCCATGGGGCTGTGTCTCTTTG 527
DB 181 CCTGGCGCCATGGGGCTGTGTCTCTTTG 210

```
RESULT 5
AAZ07518
ID AAZ07518 standard; cDNA; 177 BP.
XX
AC AAZ07518;
XX
DT 26-NOV-1999 (first entry)
XX
DE Rat HICP VWC domain encoding cDNA.
XX
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder;
KW Von Willebrand C domain; VWC.
XX
OS Rattus sp.
XX
PN W09947556-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-US05999.
XX
PR 19-MAR-1998; 98US-0044273.
XX
PA (TUFT ) TUFTS COLLEGE.
XX
PI Castello JJ;
XX
DR WPI; 1999-562060/47.
DR P-PSDB; AAY27436.
XX
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,
PT used in methods to identify modulators or in diagnostic applications
XX
PS Disclosure; Page 103; 108pp; English.
XX
CC The invention provides a rat heparin-induced CCN-like protein (HICP)
CC protein. Agents that stimulate or inhibit HICP protein activity or
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,
CC can be used to modulate cell-associated activity. HICP modulators can be
CC used to treat disorders characterized by aberrant HICP protein activity
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies
CC specific for HICP can be used to detect HICP activity in a biological
CC sample. HICP can be used to treat disorders, such as a cardiovascular or
CC fibrotic disorder, characterized by aberrant cell proliferation. The
CC present sequence represents a cDNA encoding the Von Willebrand C (VWC)
CC a domain of the HICP polypeptide.
XX
SQ Sequence 177 BP; 35 A; 47 C; 60 G; 35 T; 0 other;
XX
Query Match 14.2%; Score 90; DB 20; Length 177;
Best Local Similarity 100.0%; Pred. No. 5.5e-34;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 546 TGTGAGGTGAATGCCGCGAGGTACTGTGATGAGAGACCTTTAAACCAATTCAGGGTC 605
Db 1 TGTGAGGTGAATGCCGCGAGGTACTGTGATGAGAGACCTTTAAACCAATTCAGGGTC 60
OY 606 CTGTGCCGCTGTGATGACCGTGGCTTACC 635
Db 61 CTGTGCCGCTGTGATGACCGTGGCTTACC 90
RESULT 6
AAZ76489/C
ID AAZ76489 standard; DNA; 753 BP.
XX
AC AAZ76489;
XX
DT 06-AUG-1999 (first entry)
XX
```

```
DE Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Mus sp.
XX
PN W09921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI; 1999-337420/28.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Disclosure; Page 179-180; 284pp; English.
XX
CC The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders, including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 753 BP; 159 A; 229 C; 238 G; 127 T; 0 other;
XX
Query Match 14.2%; Score 90; DB 20; Length 753;
Best Local Similarity 100.0%; Pred. No. 5.3e-34;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 410 AGTGTGTGACGAGCGCTGGGGAGTCTGCGACCACTGCATGTCTGCGACCCAGCCA 469
Db 592 AGTGTGTGACGAGCGCTGGGGAGTCTGCGACCACTGCATGTCTGCGACCCAGCCA 533
OY 470 GGGCCTGTTTGTGACGCTGGGGCAGGCC 499
Db 532 GGGCCTGTTTGTGACGCTGGGGCAGGCC 503
RESULT 7
AAZ76488
ID AAZ76488 standard; DNA; 1734 BP.
XX
```



```
AC AAX76488;
XX
DT 06-AUG-1999 (first entry)
XX
DE Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.
XX
KM WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KM connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KM leukaemia; lymphoid malignancy; haematopoesis-related disorder;
KM tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KM kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KM connective tissue disorder; catabolic state; inflammation;
KM testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Mus sp.
XX
PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI; 1999-337420/28.
DR P-PSDB; AAY17651.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Example 2; Page 178-179; 284pp; English.
XX
CC The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;
XX
Query Match 14.2%; Score 90; DB 20; Length 1734;
Best Local Similarity 100.0%; Pred. No. 5.2e-34;
Matches .90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 410 AGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGTGCGACCCAGCCA 469
Db 418 AGTGTGCACGAGGCTGGGGAGTCTCGACACCTGCATGTGCGACCCAGCCA 477
OY 470 GGGCCTGTTTGTACGCTGGGGCAGGCC 499
Db 478 GGGCCTGTTTGTACGCTGGGGCAGGCC 507
```

```
RESULT 8
AAX76501
ID AAX76501 standard; DNA; 738 BP.
XX
AC AAX76501;
XX
DT 06-AUG-1999 (first entry)
XX
DE Human WISP-2 protein nucleotide sequence clone SEQ ID NO:38.
XX
KM WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KM connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KM leukaemia; lymphoid malignancy; haematopoesis-related disorder;
KM tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KM kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KM connective tissue disorder; catabolic state; inflammation;
KM testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Homo sapiens.
XX
PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI; 1999-337420/28.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Example 4; Page 199-200; 284pp; English.
XX
CC The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 738 BP; 104 A; 272 C; 238 G; 124 T; 0 other;
XX
Query Match 5.0%; Score 32; DB 20; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 375 GTACCCCTGTGCTGATGCTGTGCTGCTG 406
Db 115 GTACCCCTGTGCTGATGCTGTGCTGCTG 146
```



```
RESULT 9
AAx76487/c
ID AAX76487 standard; DNA; 750 BP.
XX
AC AAX76487;
XX
DT 06-AUG-1999 (first entry)
XX
DE Human WISP-2 protein complementary nucleotide sequence SEQ ID NO:14.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
OS Homo sapiens.
XX
PN WO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98WO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Penhica D, Roy MA, Wood WI;
XX
DR WPI; 1999-337420/28.
XX
PT New isolated wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Disclosure; Page 175-176; 284pp; English.
XX
CC The present invention describes wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoealic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 750 BP; 125 A; 242 C; 274 G; 109 T; 0 other;
XX
Query Match 5.0%; Score 32; DB 20; Length 750;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 375 GTACCCCTGGTGTGATGCTGTGGCTGTG 406
ID |||||
DB 624 GTACCCCTGGTGTGATGCTGTGGCTGTG 593
```

```
RESULT 10
AAx28435
ID AAX28435 standard; DNA; 1257 BP.
XX
AC AAX28435;
XX
DT 22-JUN-1999 (first entry)
XX
DE EGF-like homologue PRO261 coding sequence.
XX
KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;
KW EBAF-2; inhibitor; tumour growth; cancer; EGF-like homologue;
KW FGF-8 homologue; ss.
XX
OS Homo sapiens.
XX
PN WO9914327-A2.
XX
PD 25-MAR-1999.
XX
PF 10-SEP-1998; 98WO-US18824.
XX
PR 25-NOV-1997; 97US-0066840.
PR 17-SEP-1997; 97US-0059114.
PR 17-SEP-1997; 97US-0059117.
PR 18-SEP-1997; 97US-0059263.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 24-OCT-1997; 97US-0062816.
PR 29-OCT-1997; 97US-0063704.
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein D, Goddard A, Gurney A, Hillan K, Lawrence DA;
PI Roy M, Wood WI;
XX
DR WPI; 1999-229532/19.
DR P-PSDB; AAY05285.
XX
PT Antibodies against specific proteins overexpressed in tumours
XX
PS Example 1; Fig 22; 130pp; English.
XX
CC This sequence encodes the EGF-like homologue PRO261.
CC The invention relates to antibodies (Ab) that bind to any of the
CC polypeptides (I)-designated PRO187; PRO533; PRO214; PRO240; PRO211;
CC PRO230; PRO261; PRO246 or EBAF-2. The Ab, or other agents that inhibit
CC expression and/or activity of (I) are used: (i) to inhibit growth of
CC tumours; and (ii) as diagnostic/prognostic reagents for detection or
CC quantification of (I) in cells or tissues, by standard immunoassays, with
CC overexpression being indicative of cancer. For therapeutic use, the Ab
CC may be conjugated to a toxin, chemotherapeutic agent or radioisotope.
CC Genes expressing (I), many of which are growth factor homologues, are
CC overexpressed in some cases of cancer.
XX
SQ Sequence 1257 BP; 215 A; 416 C; 385 G; 241 T; 0 other;
XX
Query Match 5.0%; Score 32; DB 20; Length 1257;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 11
AAA30048
ID AAA30048 standard; cDNA; 1266 BP.
XX
AC AAA30048;
```

XX 09-AUG-2000 (first entry)
XX
DE Human PRO261 nucleotide sequence.
XX
KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;
KW PRO317; tumour growth inhibitor; cancer; diagnosis; treatment; human;
KW cell growth; proliferation; growth factor; ADEPT;
KW antibody dependent enzyme mediated prodrug therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200015666-A2.
XX
PD 23-MAR-2000.
XX
PF 08-SEP-1999; 99WO-US20594.
XX
PR 10-SEP-1998; 98US-0099803.
PR 10-SEP-1998; 98WO-US18824.
XX
PA (GETH) GENENTECH INC.
XX
PI Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI, Botstein D;
XX
DR WPI; 2000-271386/23.
DR P-PSDB; AAY88573.
XX
XX
PT New isolated antibodies which bind to specific polypeptides used for
PT diagnosis and treatment of neoplastic cell growth and proliferation -
XX
XX
PS Example 7; Fig 13; 200pp; English.
XX
XX This sequence represents a human PRO261 nucleotide sequence. PRO261 is a
CC growth factor. The invention relates to isolated antibodies which bind to
CC a polypeptide. The "PRO" polypeptides are encoded by genes which are over
CC expressed in the genome of tumour cells. Vectors and host cells
CC comprising the nucleic acid encoding the antibodies are used in the
CC production of the antibodies. The antibodies and nucleic acids encoding
CC them are used for diagnosing a tumour in a mammal. The antibodies are
CC used for inhibiting the growth of tumour cells and identifying compounds
CC that inhibit a biological or immunological activity of and/or expression
CC of a PRO187, PRO533, PRO214, PRO240, PRO211, PRO230, PRO261, PRO246 or
CC PRO317 polypeptide. The antibody can be used in antibody dependent enzyme
CC mediated prodrug therapy (ADEPT) by conjugating the antibody to a
CC prodrug-activating enzyme which converts a prodrug to an anti-cancer
CC drug. The antibodies can be fluorescently labelled and monitored by light
CC microscopy, flow cytometry or fluorimetry for diagnosis and prognosis of
CC tumours.
XX
SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;
XX
XX
Query Match 5.0%; Score 32; DB 21; Length 1266;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
|||
Db 136 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 167
XX
RESULT 12
AAS21403
ID AAS21403 standard; cDNA; 1266 BP.
XX
AC AAS21403;
XX
DT 24-OCT-2001 (first entry)
XX
XX Human cDNA sequence encoding for PRO261 polypeptide.
DE
XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW

KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200140466-A2.
XX
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000WO-US32678.
XX
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.
XX
XX
PA (GETH) GENENTECH INC.
XX
XX
PI Baker KP, Beresini M, Deforme L, Desnoyers L, Flivaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-408281/43.
DR P-PSDB; AAU12331.
XX
XX
PT Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
XX lung, breast, prostate, cervical
XX
PS Claim 3; Fig 319; 813pp; English.
XX
XX
AAS21244-AAS21518 encode for novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIA. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding

CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.

SO Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 5.0%; Score 32; DB 22; Length 1266;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 GTACCCCTGTGCTGATGGCTGTGCTGCTG 406
Db 136 GTACCCCTGTGCTGATGGCTGTGCTGCTG 167

RESULT 13
AAAF60368
ID AAF60368 standard; cDNA; 1266 BP.

AC AAF60368;
DT 27-APR-2001 (first entry)

DE PRO261 coding sequence.

KW Cytostatic; PRO protein; tumour; cancer; ss.

OS Homo sapiens.

PN WO200105836-A1.

PD 25-JAN-2001.

PF 20-DEC-1999; 99WO-US30999.

PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28564.

PA (GETH) GENENTECH INC.

PI Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;

DR WPI; 2001-091968/10.

DR P-PSDB; AAB68598.

PT New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,
PT useful for diagnosing and treating cancers -

PS Claim 50; Fig 13; 196pp; English.

CC The present invention relates to PRO proteins and coding sequences. The
CC present sequence is the coding sequence for one such PRO protein.
CC It was found that the PRO genes are amplified in the genome of tumour
CC cells. The gene amplification is expected to be associated with the
CC overexpression of the gene product and contributes to tumourigenesis.
CC Therefore, antagonists of PRO proteins are useful for the treatment of
CC benign or malignant tumours, leukaemias, lymphoid malignancies and other
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
CC epithelial, inflammatory and immunologic disorders.

SO Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 5.0%; Score 32; DB 22; Length 1266;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 GTACCCCTGTGCTGATGGCTGTGCTGCTG 406

Db 136 GTACCCCTGTGCTGATGGCTGTGCTGCTG 167

RESULT 14
AAC97451
ID AAC97451 standard; cDNA; 1266 BP.

AC AAC97451;

DT 28-FEB-2001 (first entry)

DE Human angiogenesis-associated protein PRO261 cDNA, SEQ ID NO:107.

KW Human; angiogenesis-associated protein; PRO; endothelial cell growth;
KW cardiac hypertrophy; cardiovascular disorder; endothelial disorder;
KW angiogenic disorder; atherosclerosis; osteoporosis; hypertension;
KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;
KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
KW gene therapy; transgenic animal; ss.

OS Homo sapiens.

PN WO200053753-A2.

PD 14-SEP-2000.

PF 05-JAN-2000; 2000WO-US00219.

PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.

PA (GETH) GENENTECH INC.

PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;

PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;

PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;

DR WPI; 2001-090793/10.
DR P-PSDB; AAB53084.

PT New isolated nucleic acid for producing a PRO polypeptide, analyzing
PT genetic disorders and treating cardiovascular, endothelial or
PT angiogenic disorders, such as atherosclerosis, wounds or cancer -
PS Claim 58; Fig 41; 293pp; English.

CC The invention relates to novel human angiogenesis-associated proteins
CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding
CC PRO proteins. The invention also relates to vectors and host cells
CC comprising a PRO nucleic acid, the recombinant production of a PRO
CC protein, PRO antibodies specific for a PRO protein, fusion proteins
CC comprising a PRO protein, agonists or antagonists of a PRO protein, and
CC compounds which inhibit the expression of a PRO gene. The invention
CC additionally encompasses methods of identifying modulators of PRO
CC expression or activity; diagnosing a cardiovascular, endothelial or
CC angiogenic disorder, or a susceptibility to such a disorder by detecting
CC mutations in a PRO gene, or the expression level of a PRO gene within a

```
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 18
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-182-145-18
```

```
Query Match      14.2%; Score 90; DB 4; Length 1734;
Best Local Similarity 100.0%; Pred. No. 4.2e-35;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      410 AGTGTGTCACGAGGCGCTGGGGAGTCTCTGCAGACCACCTGCATGTCGCGAGCCAGCCCA 469
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1317 AGTGTGTCACGAGGCGCTGGGGAGTCTCTGCAGACCACCTGCATGTCGCGAGCCAGCCCA 1258

QY      470 GGGCCTGGTTGTCTCAGCCTGGGGCAGGCC 499
      |||||||||||||||||||||||||||||||
Db      1257 GGGCCTGGTTGTCTCAGCCTGGGGCAGGCC 1228
```

```
RESULT 3
US-09-182-145-38
; Sequence 38, Application US/09182145B
; Patent No. 6387657
```

GENERAL INFORMATION:

```
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 38
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-182-145-38
```

```
Query Match      5.0%; Score 32; DB 4; Length 738;
```

```
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
      |||||||||||||||||||||||||||||||
Db      115 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 146
```

RESULT 4

```
US-09-182-145-13
; Sequence 13, Application US/09182145B
; Patent No. 6387657
```

GENERAL INFORMATION:

```
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 13
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-182-145-13
```

```
Query Match      5.0%; Score 32; DB 4; Length 1293;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      375 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 406
      |||||||||||||||||||||||||||||||
Db      148 GTACCCCTGCTGCTGATGGCTGTGGCTGCTG 179
```

RESULT 5

```
US-09-182-145-14/C
; Sequence 14, Application US/09182145B
; Patent No. 6387657
```

GENERAL INFORMATION:

```
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:50:06 ; Search time 34.7003 Seconds
(without alignments)
5612.050 Million cell updates/sec

Title: US-10-010-408-1_COPY_1_635
Perfect score: 635
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:**
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:**
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:**
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:**
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:**
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	14.2	1734	4 US-09-182-145-17	Sequence 17, Appl
2	90	14.2	1734	4 US-09-182-145-18	Sequence 18, Appl
3	32	5.0	1738	4 US-09-182-145-38	Sequence 38, Appl
4	32	5.0	1293	4 US-09-182-145-13	Sequence 13, Appl
5	32	5.0	1293	4 US-09-182-145-14	Sequence 14, Appl
6	27	4.3	51	4 US-09-182-145-117	Sequence 117, App
7	19	3.0	616	4 US-09-385-982-220	Sequence 220, App
8	19	3.0	841	4 US-09-182-145-39	Sequence 39, Appl
9	19	3.0	1196	4 US-09-149-476-225	Sequence 225, App
10	19	3.0	1220	4 US-09-149-476-57	Sequence 57, Appl
11	19	3.0	1514	4 US-09-213-768-1	Sequence 1, Appli
12	19	3.0	1539	4 US-09-668-680-13	Sequence 13, Appl
13	18	2.8	20	2 US-09-213-768-2	Sequence 2, Appli
14	18	2.8	2949	4 US-08-259-451-10	Sequence 10, Appl
15	18	2.8	3727	1 US-08-249-380-1	Sequence 1, Appli
16	18	2.8	8957	4 US-08-259-451-1	Sequence 1, Appli
17	17	2.7	44	4 US-09-182-145-152	Sequence 152, App
18	17	2.7	480	3 US-09-188-930-206	Sequence 206, App
19	17	2.7	614	4 US-08-998-416-151	Sequence 151, App
20	17	2.7	661	3 US-08-578-634C-3	Sequence 3, Appli
21	17	2.7	661	4 US-09-430-010-3	Sequence 3, Appli
22	17	2.7	896	3 US-09-188-930-36	Sequence 36, Appl
23	17	2.7	1491	4 US-09-662-249A-3	Sequence 3, Appli
24	17	2.7	1743	3 US-08-665-259-20	Sequence 20, Appl
25	17	2.7	1743	3 US-08-762-500-20	Sequence 20, Appl
26	17	2.7	1974	3 US-08-762-500-78	Sequence 78, Appl
27	17	2.7	3280	4 US-09-484-970B-43	Sequence 43, Appl

28	17	2.7	3865	4 US-09-149-476-296	Sequence 296, App
29	17	2.7	6803	3 US-08-665-259-19	Sequence 19, Appl
30	17	2.7	6803	3 US-08-762-500-19	Sequence 19, Appl
31	17	2.7	7874	4 US-09-780-175-96	Sequence 96, Appl
32	17	2.7	35828	4 US-09-449-218D-17	Sequence 17, Appl
33	17	2.7	81001	4 US-09-750-580-1	Sequence 1, Appli
34	17	2.7	111282	4 US-09-754-250-3	Sequence 3, Appli
35	16	2.5	428	4 US-09-397-787-224	Sequence 224, App
36	16	2.5	435	4 US-09-397-787-280	Sequence 280, App
37	16	2.5	572	4 US-09-342-653-5	Sequence 5, Appli
38	16	2.5	691	1 US-08-243-879A-37	Sequence 37, Appl
39	16	2.5	691	1 US-08-499-523-3	Sequence 3, Appli
40	16	2.5	691	3 US-09-128-345-3	Sequence 3, Appli
41	16	2.5	760	3 US-08-768-859A-7	Sequence 7, Appli
42	16	2.5	760	3 US-08-767-820A-7	Sequence 7, Appli
43	16	2.5	760	5 PCT-US95-06157-7	Sequence 7, Appli
44	16	2.5	766	3 US-08-768-859A-9	Sequence 9, Appli
45	16	2.5	766	3 US-08-767-820A-9	Sequence 9, Appli

ALIGNMENTS

```
RESULT 1
US-09-182-145-17
; Sequence 17, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-182-145-17

Query Match      14.2%  Score 90;  DB 4;  Length 1734;
Best Local Similarity 100.0%;  Pred. No. 4.2e-35;
Matches 90;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

OY      410 AGTGTGTCACGAGAGGCTGGGGAGTCTCGACACCACTGCATGTCTGGACCCAGCCA 469
          |||||||
Db      418 AGTGTGTCACGAGAGGCTGGGGAGTCTCGACACCACTGCATGTCTGGACCCAGCCA 477

OY      470 GGGCCTGTTGTTCAGCCTGGGGCAGGCC 499
          |||||||
Db      478 GGGCCTGTTGTTCAGCCTGGGGCAGGCC 507

RESULT 2
US-09-182-145-18/c
; Sequence 18, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
```

The present invention describes nucleic acids encoding PRO polypeptides useful for preventing, diagnosing and treating diagnosing a cardiovascular, endothelial or angiogenic disorder in mammals by modulating cell proliferation, angiogenesis and cardiovascularisation,

.....

THIS PAGE BLANK (USPTO)